

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 12:17:57 ; Search time 6270.82 Seconds
(without alignments)
11955.722 Million cell updates/sec

Title: US-09-528-031-1

Perfect score: 4847
Sequence: 1 GCGTCATGCTCGGAGCGTG.....AAAAAAAAAGGGCGGCCGC 4847

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_om: *
8: gb_ov: *
9: gb_pat1: *
10: gb_pat2: *
11: gb_ph: *
12: gb_pl1: *
13: gb_pl2: *
14: gb_pl3: *
15: gb_pl4: *
16: em_ba1: *
17: em_ba2: *
18: em_fun: *
19: em_htgo_hum: *
20: em_htgo_inv: *
21: em_htgo_rtd: *
22: em_htg_hum1: *
23: em_htg_hum2: *
24: em_htg_hum3: *
25: em_htg_hum4: *
26: em_htg_hum5: *
27: em_htg_hum6: *
28: em_htg_hum7: *
29: em_htg_hum8: *
30: em_htg_inv1: *
31: em_htg_inv2: *
32: em_htg_other: *
33: em_htg_rtd: *
34: em_hum1: *
35: em_hum2: *
36: em_hum3: *
37: em_hum4: *
38: em_hum5: *
39: em_hum6: *
40: em_hum7: *
41: em_hum8: *
42: em_om: *
43: em_or: *

44: em_ov: *
45: em_pat: *
46: em_ph: *
47: em_pl: *
48: em_ro: *
49: em_sts: *
50: em_sy: *
51: em_un: *
52: em_v1: *
53: gb_sts1: *
54: gb_sts2: *
55: gb_sts3: *
56: gb_sy: *
57: gb_un: *
58: gb_v1: *
59: gb_v12: *
60: gb_htg1: *
61: gb_htg2: *
62: gb_htg3: *
63: gb_htg4: *
64: gb_htg5: *
65: gb_htg6: *
66: gb_htg7: *
67: gb_htg8: *
68: gb_htg9: *
69: gb_htg10: *
70: gb_htg11: *
71: gb_htg12: *
72: gb_htg13: *
73: gb_htg14: *
74: gb_htg15: *
75: gb_htg16: *
76: gb_htg17: *
77: gb_htg18: *
78: gb_htg19: *
79: gb_htg20: *
80: gb_htg21: *
81: gb_htg22: *
82: gb_htg23: *
83: gb_htg24: *
84: gb_htg25: *
85: gb_pr1: *
86: gb_pr2: *
87: gb_pr3: *
88: gb_pr4: *
89: gb_pr5: *
90: gb_pr6: *
91: gb_pr7: *
92: gb_pr8: *
93: gb_pr9: *
94: gb_r01: *
95: gb_r02: *
96: gb_in4: *
97: gb_pr10: *
98: em_ba3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	4847	100.0	4847	9	AR099619 Sequence
2	4792.6	98.9	5838	88	AF104942 Homo sapi
3	4789.6	98.8	5881	89	AF146074 Homo sapi
4	4788	98.8	5826	97	HS083661 Homo sapien
5	4775.4	98.5	4781	9	AR091520 Sequence
6	4728	97.5	5728	85	AB019002 Homo sapi
7	3774	77.9	5400	94	AB019003 Mus muscu
8	3703.8	76.4	4939	85	AB005659 Homo sapi

9	3691.4	76.2	5744	94	AB020209	AB020209 Rattus no
10	1674.4	34.5	1761	9	A64695	A64695 Sequence 61
11	1054.2	21.7	2167	9	A64696	A64696 Sequence 62
12	776	16.0	4492	89	AF352582	AF352582 Homo sapi
13	545	11.2	765	94	AB012090	AB012090 Mus muscu
14	534.2	11.0	1076	94	AF213387	AF213387 Mus muscu
15	511.2	10.5	5011	9	AR070199	AR070199 Sequence
16	511.2	10.5	5011	9	AR070199	AR070199 Sequence
17	511.2	10.5	5011	97	H08MRP	H08MRP Human multi
18	508	10.5	5867	97	SATMRP	SATMRP Human multi
19	506.4	10.4	5011	9	AR070198	AR070198 Sequence
20	506.4	10.4	5011	9	AR093439	AR093439 Sequence
21	506.4	10.4	5011	10	I17455	I17455 Sequence 1
22	479.8	9.9	188429	73	AC068644	AC068644 Homo sapi
23	437.4	9.0	2940	88	AF071203	AF071203 Homo sapi
24	437.4	9.0	4231	88	AF071203	AF071203 Homo sapi
25	420.4	8.7	4587	94	AF022908	AF022908 Mus muscu
26	420.4	8.7	5889	9	AR070200	AR070200 Sequence
27	420.4	8.7	5889	9	AR093441	AR093441 Sequence
28	406.6	8.4	2969	95	RN0277881	RN0277881 Rattus no
29	390.4	8.1	4880	94	D86086	D86086 Rat mRNA fo
30	390.4	8.1	4918	9	A64636	A64636 Sequence 2
31	390.4	8.1	4918	95	RATCMOAT	RATCMOAT Rattus norv
32	388.8	8.0	4909	95	RNCMRP	RNCMRP Rattus norv
33	385.4	8.0	5023	89	AK000002	AK000002 Homo sapi
34	384	7.9	5407	7	OC249144	OC249144 O. cuniculus
35	383.8	7.9	4868	97	HS063970	HS063970 Human canal
36	382.2	7.9	5300	97	HS049248	HS049248 Human canal
37	382.2	7.9	5586	9	A64635	A64635 Sequence 1
38	380.6	7.9	4864	93	HSCMRP	HSCMRP H. sapiens m
39	380.6	7.9	4868	10	EL5807	EL5807 Human mRNA
40	376.8	7.8	4749	95	RNSULFUR	RNSULFUR R. norvegicu
41	376.8	7.8	4780	94	AB052294	AB052294 Rattus no
42	376.8	7.8	5099	95	RAT5UR	RAT5UR Rattus norv
43	375.2	7.7	166918	4	AC009205	AC009205 Drosophila
44	375.2	7.7	170546	4	AC011756	AC011756 Drosophila
45	375.2	7.7	265010	5	AE003661	AE003661 Drosophila

ALIGNMENTS

RESULT	1	4847 bp	DNA	PAT	14-FEB-2001
AR099619	Sequence 1	from patent US 6077936.			
LOCUS	AR099619				
DEFINITION	Sequence 1	from patent US 6077936.			
ACCESSION	AR099619				
VERSION	AR099619.1	GI:12809385			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 4847)				
AUTHORS	Shyjan, A.				
TITLE	Multidrug resistance-associated polypeptide				
JOURNAL	Patent: US 6077936-A 1 20-JUN-2000;				
FEATURES	Location/Qualifiers				
source	1..4847				
BASE COUNT	1226 a 1178 c 1263 g 1180 t				
ORIGIN					

Query Match 100.0%; Score 4847; DB 9; Length 4847;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4847; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGCTCATGCTCGGGAGCGGTGTGAGCGGCTGCTCGAGCAGGGGCGGAG	60
DB	1	GGCTCATGCTCGGGAGCGGTGTGAGCGGCTGCTCGAGCAGGGGCGGAG	60
QY	61	GAATTCGATGTAACACTACAGTCTGTGAGCCTTGAACTCCACTCAGAGAAATGAA	120
DB	61	GAATTCGATGTAACACTACAGTCTGTGAGCCTTGAACTCCACTCAGAGAAATGAA	120

1201 TACTACATTAATTATCAAAATGATGCCGGGCAAAAGCATTTTCACAGATGTCA 1260
1201 TACTACATTAATTATCAAAATGATGCCGGGCAAAAGCATTTTCACAGATGTCA 1260
1261 GAAATCCGCGAGGAGGAGCGTGGATATTGGAAAAAGCGGGTACTTCCAGAGCATCAC 1320
1261 GAAATCCGCGAGGAGGAGCGTGGATATTGGAAAAAGCGGGTACTTCCAGAGCATCAC 1320
1321 TGTGGGTGGGTCCCATTTGTGGTGTATTCACAGCGTGGTACCTTCTGTTCATAT 1380
1321 TGTGGGTGGGTCCCATTTGTGGTGTATTCACAGCGTGGTACCTTCTGTTCATAT 1380
1321 TGTGGGTGGGTCCCATTTGTGGTGTATTCACAGCGTGGTACCTTCTGTTCATAT 1380
1381 GACCGTGGGCTTCAGATTCAGACAGACAGAGCTTTCACAGTGGTACAGTCTTCATAT 1440
1381 GACCGTGGGCTTCAGATTCAGACAGACAGAGCTTTCACAGTGGTACAGTCTTCATAT 1440
1381 GACCGTGGGCTTCAGATTCAGACAGACAGAGCTTTCACAGTGGTACAGTCTTCATAT 1440
1441 CATGACCTTTTGTGAAAGTAACACCGTTTTCAGTAAAGTCCCTCTCAGAAAGCTCAGT 1500
1441 CATGACCTTTTGTGAAAGTAACACCGTTTTCAGTAAAGTCCCTCTCAGAAAGCTCAGT 1500
1501 GGCTGTGACATTTAAGAGTTGTCTTAATGGAAGAGTTTACATATTAAGACAA 1560
1501 GGCTGTGACATTTAAGAGTTGTCTTAATGGAAGAGTTTACATATTAAGACAA 1560
1501 GGCTGTGACATTTAAGAGTTGTCTTAATGGAAGAGTTTACATATTAAGACAA 1560
1561 ACCAGCAGATCCTCAGATCAGATGAGATGAAAAATGCCACCTTGGCATGGAGCTCTC 1620
1561 ACCAGCAGATCCTCAGATCAGATGAGATGAAAAATGCCACCTTGGCATGGAGCTCTC 1620
1561 ACCAGCAGATCCTCAGATCAGATGAGATGAAAAATGCCACCTTGGCATGGAGCTCTC 1620
1621 CCACCTCAGATTCAGAAAGTCCGCCAAGCTGACCCCAAAATGAAAAAGACAGAGGCG 1680
1621 CCACCTCAGATTCAGAAAGTCCGCCAAGCTGACCCCAAAATGAAAAAGACAGAGGCG 1680
1681 TTCAGAGGCGCAAGAGAGAGAGTGAAGCTGACAGCGCTGACATTCAGAGCGGTGCT 1740
1681 TTCAGAGGCGCAAGAGAGAGAGTGAAGCTGACAGCGCTGACATTCAGAGCGGTGCT 1740
1681 TTCAGAGGCGCAAGAGAGAGAGTGAAGCTGACAGCGCTGACATTCAGAGCGGTGCT 1740
1741 GGCAGAGCGAAAGAGCGACCTCTCTGAGACGTGACAGCGCGCCAGTCCGAAAGAGA 1800
1741 GGCAGAGCGAAAGAGCGACCTCTCTGAGACGTGACAGCGCGCCAGTCCGAAAGAGA 1800
1741 GGCAGAGCGAAAGAGCGACCTCTCTGAGACGTGACAGCGCGCCAGTCCGAAAGAGA 1800
1801 AGAAGGCAAGCAGATCCACCTGGGCACTGCGCTTACAGAGGACACTGACAGCATCGA 1860
1801 AGAAGGCAAGCAGATCCACCTGGGCACTGCGCTTACAGAGGACACTGACAGCATCGA 1860
1801 AGAAGGCAAGCAGATCCACCTGGGCACTGCGCTTACAGAGGACACTGACAGCATCGA 1860
1861 TCTGAGATCCAGAGGCTAACTGTTGGAATCTGCGCAGTGTGGAAATGGAAGAAAC 1920
1861 TCTGAGATCCAGAGGCTAACTGTTGGAATCTGCGCAGTGTGGAAATGGAAGAAAC 1920
1861 TCTGAGATCCAGAGGCTAACTGTTGGAATCTGCGCAGTGTGGAAATGGAAGAAAC 1920
1921 CTTCTCATTTAGGCAATTTAGGCGAGATGACGCTTTCAGAGGCGAGATTCAGATCAG 1980
1921 CTTCTCATTTAGGCAATTTAGGCGAGATGACGCTTTCAGAGGCGAGATTCAGATCAG 1980
1921 CTTCTCATTTAGGCAATTTAGGCGAGATGACGCTTTCAGAGGCGAGATTCAGATCAG 1980
1981 TGAAGCCTTCGCTTATGTGGCCAGAGCGCTGATCTCAATGTACTCTGAGAGACAA 2040
1981 TGAAGCCTTCGCTTATGTGGCCAGAGCGCTGATCTCAATGTACTCTGAGAGACAA 2040
1981 TGAAGCCTTCGCTTATGTGGCCAGAGCGCTGATCTCAATGTACTCTGAGAGACAA 2040
2041 CATCTGTTTGGGAAGAAATGATGAAAGAAATCAACTCTGTGCTGACAGCTGCTG 2100
2041 CATCTGTTTGGGAAGAAATGATGAAAGAAATCAACTCTGTGCTGACAGCTGCTG 2100
2041 CATCTGTTTGGGAAGAAATGATGAAAGAAATCAACTCTGTGCTGACAGCTGCTG 2100
2101 CCGTGGAGGCTGACCTGGCCATTTCTCCAGACGACCTGACAGGAAATGGAGAGGAG 2160
2101 CCGTGGAGGCTGACCTGGCCATTTCTCCAGACGACCTGACAGGAAATGGAGAGGAG 2160
2101 CCGTGGAGGCTGACCTGGCCATTTCTCCAGACGACCTGACAGGAAATGGAGAGGAG 2160
2161 AGCAACCTGACGCTGGGCGACGCGCAGAGATCAGCTTGGCCGGGCTTGTATATGTA 2220
2161 AGCAACCTGACGCTGGGCGACGCGCAGAGATCAGCTTGGCCGGGCTTGTATATGTA 2220
2161 AGCAACCTGACGCTGGGCGACGCGCAGAGATCAGCTTGGCCGGGCTTGTATATGTA 2220
2221 CAGGAGCATCTACATCTGAGACACCCCTCAGTGCCTTATGATGCCATGTGGCAACCA 2280
2221 CAGGAGCATCTACATCTGAGACACCCCTCAGTGCCTTATGATGCCATGTGGCAACCA 2280

2281 CATCTCAATAGTGTATCCGGAACATCTCAAGTCCAGAGATCTGTGTTACCA 2340
2281 CATCTCAATAGTGTATCCGGAACATCTCAAGTCCAGAGATCTGTGTTACCA 2340
2341 CAGTGTACATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
2341 CAGTGTACATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
2401 GGAAGAGGCGACCATGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
2401 GGAAGAGGCGACCATGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
2461 TAACTGTTGCTGGGAGAGACACCGCATGATGATGATGATGATGATGATGATGATGATGAT 2520
2461 TAACTGTTGCTGGGAGAGACACCGCATGATGATGATGATGATGATGATGATGATGATGAT 2520
2461 TAACTGTTGCTGGGAGAGACACCGCATGATGATGATGATGATGATGATGATGATGATGAT 2520
2521 TTCACAGAAAGTCAACAGACAGAGGCTTAAACAGATCAATTAAGAGAGAAAGC 2580
2521 TTCACAGAAAGTCAACAGACAGAGGCTTAAACAGATCAATTAAGAGAGAAAGC 2580
2521 TTCACAGAAAGTCAACAGACAGAGGCTTAAACAGATCAATTAAGAGAGAAAGC 2580
2581 AGTAAAGCCAGAGAGAGGCGACCTGTGACGTGGAGAGAAAGGCGAGGTTCAAGTGC 2640
2581 AGTAAAGCCAGAGAGAGGCGACCTGTGACGTGGAGAGAAAGGCGAGGTTCAAGTGC 2640
2581 AGTAAAGCCAGAGAGAGGCGACCTGTGACGTGGAGAGAAAGGCGAGGTTCAAGTGC 2640
2641 CTGCTCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
2641 CTGCTCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
2701 TATGCGCTTTTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
2701 TATGCGCTTTTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
2701 TATGCGCTTTTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
2761 CTGATCAAGAGAGAGAGGAGAGACACCTGATGATGATGATGATGATGATGATGATGATGAT 2820
2761 CTGATCAAGAGAGAGAGGAGAGACACCTGATGATGATGATGATGATGATGATGATGATGAT 2820
2761 CTGATCAAGAGAGAGAGGAGAGACACCTGATGATGATGATGATGATGATGATGATGATGAT 2820
2821 TGAAGCATGAGAGAGACAACTCATATGATGATGATGATGATGATGATGATGATGATGAT 2880
2821 TGAAGCATGAGAGAGACAACTCATATGATGATGATGATGATGATGATGATGATGATGAT 2880
2821 TGAAGCATGAGAGAGACAACTCATATGATGATGATGATGATGATGATGATGATGATGAT 2880
2881 GGCAGTATGCTGATCTGAAAGCCATTGAGAGAGTGTCTTGTCAAGGACGCTGCG 2940
2881 GGCAGTATGCTGATCTGAAAGCCATTGAGAGAGTGTCTTGTCAAGGACGCTGCG 2940
2881 GGCAGTATGCTGATCTGAAAGCCATTGAGAGAGTGTCTTGTCAAGGACGCTGCG 2940
2941 AGCTTCTCCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
2941 AGCTTCTCCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
2941 AGCTTCTCCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
3001 TTTTGCAGAGACCCCAAGAGGAGATCTCAACAGGTTTCCAAAGACATGATGATGAT 3060
3001 TTTTGCAGAGACCCCAAGAGGAGATCTCAACAGGTTTCCAAAGACATGATGATGAT 3060
3001 TTTTGCAGAGACCCCAAGAGGAGATCTCAACAGGTTTCCAAAGACATGATGATGAT 3060
3061 TGAAGTGGGCTGCGCTTCCAGGCGAGATGATGATGATGATGATGATGATGATGATGATGAT 3120
3061 TGAAGTGGGCTGCGCTTCCAGGCGAGATGATGATGATGATGATGATGATGATGATGATGAT 3120
3121 CTGTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
3121 CTGTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
3121 CTGTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
3181 CATCTCTTTTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
3181 CATCTCTTTTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
3181 CATCTCTTTTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
3241 GGACATATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
3241 GGACATATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
3241 GGACATATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
3301 CACATCCAGCGCTTACATTAAGAGAGAGTCTGACAGATACAGAGACCTGTGGA 3360
3301 CACATCCAGCGCTTACATTAAGAGAGAGTCTGACAGATACAGAGACCTGTGGA 3360
3301 CACATCCAGCGCTTACATTAAGAGAGAGTCTGACAGATACAGAGACCTGTGGA 3360
3361 TGAACAACAAAGCTCTTTTGTGTTAGTGTGAGTGGGTGCTGCTGTGCGGCT 3420
3361 TGAACAACAAAGCTCTTTTGTGTTAGTGTGAGTGGGTGCTGCTGTGCGGCT 3420

SPKLPKMKDKRKRASRGKKEKKEVROJORTHOAVLAEOKHLLDSDERSPEEBKHI
IHGHLRLQRTLSHLDLEIOBKLVLGJGSGKTSKTSLSIILQMTLLBSSTASCT
FAYVAAQAMITLNLATLRLNITRKEDERINSVLNSCCLRPDLILPESDLEJTEGERS
ANLSGGQORISLALYSRISYILDDPSALAHVNHFNPAIRHLKSKVLEFV
THOLOYLVDCDEVIFMRKEGCIYTERGTHELNLUNGEDVATITNNLLGCTPEVLENSK
ETSSOKSODRQKTSVKKEKAVPEEGQLOVEEYSGVYIOAAGP
LAFIVMALFPLANYSTAFSTWMLSYMKGSGNTYTRGNETSOSDKMNPWOY
ASIIYALSMVWLIIKAIKRGVYFVKTILBASRLDELFRILRSPMKFDDTPTRILL
NRESKDEVDVRLPFQEMFIQNVILVFCVGMAGVFLVAVGLVLEFVSLH
VSRVILRELKRLDITQSPFLSHITSIOGLATIHANXGOEFLHRYOELLDDNOAF
FLFCAMKMLVRLDLISALITTTGLMIVLMHOIPEYVAGLISYAVOLTLGFOFT
VRLASEEARPTSEVERINHYIKTSLSEPAIRIKKAPSPDSEBEYFENAEKMYRE
NLPLVLEKVSFTIKPREKIGIVRGSGSKSLGMAPFLVLESGCIRIDGVNRSDIG
LADRSKLSIIRPQVLEFSVRNLDPFNQNTEDQIMDALETHMKECINOLPKLE
SEVENMDNEFSVGROLLICIRALIRKILILDEATRAMETFDLILQETIRAPAD
CTMLTIAHRLHTVJGSDRIWLAQOVVEFDPVLSLNSDSREFYAFNAENKAVAK
G⁹

BASE COUNT 1422 a 1432 c 1509 g 1475 t
ORIGIN

Query Match 98.9% Score 4792.6; DB 88; Length 5838;
Best Local Similarity 99.8%; Pred. No. 0;

Matches 4809; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

OY 1 GGCTATGCTGGAGCGGTGTGAGCGGCTGGCGGTTGTCTGAGAGAGGGGCGAG 60
DB 11 GGCTATGCTGGAGCGGTGTGAGCGGCTGGCGGTTGTCTGAGAGAGGGGCGAG 70
OY 61 GAATTCGATGTAAGTAACAGTCTGTGAGCCCTGGAACCTCCATCAGAGAAGTGA 120
DB 71 GAATTCGATGTAAGTAACAGTCTGTGAGCCCTGGAACCTCCATCAGAGAAGTGA 130
OY 121 GGATATCGACATAGAGAAAGAGTATATCATCCCACTCTGGGTATAGAGTGAAGGA 180
DB 131 GGATATCGACATAGAGAAAGAGTATATCATCCCACTCTGGGTATAGAGTGAAGGA 190
OY 181 GGAACACGACACTTCTGGGAGCAGACAGACCGTGAAGATTCCAGTTGAGAGAACTCG 240
DB 191 GGAACACGACACTTCTGGGAGCAGACAGACCGTGAAGATTCCAGTTGAGAGAACTCG 250
OY 241 ACCGTTGGAATCCAGAGATGCTTGGAAACAGACGCCGAGGGGCTCTCTCTTGA 300
DB 251 ACCGTTGGAATCCAGAGATGCTTGGAAACAGACGCCGAGGGGCTCTCTCTTGA 310
OY 301 TCCCTCCATGCAATTTCTCAGACTCAGAAATCTGATGAGAGCAATCCCAAGGAAATGACA 360
DB 311 TCCCTCCATGCAATTTCTCAGACTCAGAAATCTGATGAGAGCAATCCCAAGGAAATGACA 370
OY 361 TCATGAGCTGAGTGCCTGAGACCCATCCGAGCTACTTCAAAACACAGACCCAGTGA 420
DB 371 TCATGAGCTGAGTGCCTGAGACCCATCCGAGCTACTTCAAAACACAGACCCAGTGA 430
OY 421 CAATGCTGGGCTTTTTCCTGTATGACTTTTCGTGGCTTTCCTCTGTCGCCCTGTGTGC 480
DB 431 CAATGCTGGGCTTTTTCCTGTATGACTTTTCGTGGCTTTCCTCTGTCGCCCTGTGTGC 490
OY 481 CCACAGAAGAGGGGAGCTTCAATGAGAGAGTGTGTCTGTCTCAACAGAGTCTTC 540
DB 491 CCACAGAAGAGGGGAGCTTCAATGAGAGAGTGTGTCTGTCTCAACAGAGTCTTC 550
OY 541 TCACGCGAAGCTCAGAGAAGCTAGAGAGTGTGGCAAGAGAGCTGAAGAGTTGGGCC 600
DB 551 TCACGCGAAGCTCAGAGAAGCTAGAGAGTGTGGCAAGAGAGCTGAAGAGTTGGGCC 610
OY 601 AACAGCTGCTTCCCTGCGAAGGTTGTGATCTTCTCCACACAGAGCTATCTCTGTC 660
DB 611 AACAGCTGCTTCCCTGCGAAGGTTGTGATCTTCTCCACACAGAGCTATCTCTGTC 670
OY 661 CATGCTGTGCTGATGATACGACAGCTGGCTGGCTTCACTGAGACCAAGCTTCAATGTTGA 720
DB 671 CATGCTGTGCTGATGATACGACAGCTGGCTGGCTTCACTGAGACCAAGCTTCAATGTTGA 730

OY 721 ACACCTTTGAGATATACCCAGGCAACAGAGCTAATCCGTCAGTACAGCTTGTGTAGT 780
DB 731 ACACCTTTGAGATATACCCAGGCAACAGAGCTAATCCGTCAGTACAGCTTGTGTAGT 790
OY 781 GCTGGGCTTCCCTCCAGAGGAATGTCGGGCTTGTGCTGCTTGCATGACTTGGGCAAT 840
DB 791 GCTGGGCTTCCCTCCAGAGGAATGTCGGGCTTGTGCTGCTTGCATGACTTGGGCAAT 850
OY 841 GAATTAACGAAACCGGTGCTCCCTTCGGGGGGCCATCTCAACATGCAATTAAGAGAT 900
DB 851 GAATTAACGAAACCGGTGCTCCCTTCGGGGGGCCATCTCAACATGCAATTAAGAGAT 910
OY 901 CCTTAAGTTAAAGAACATTAAGAGAAATCCCTGGGTGAGCTCATCAATTTGCTCAA 960
DB 911 CCTTAAGTTAAAGAACATTAAGAGAAATCCCTGGGTGAGCTCATCAATTTGCTCAA 970
OY 961 CGATGGGCAAGAAATGTTTGAAGCAGACAGCCGTTGGACAGCTGTGCTGAGAGACCGT 1020
DB 971 CGATGGGCAAGAAATGTTTGAAGCAGACAGCCGTTGGACAGCTGTGCTGAGAGACCGT 1030
OY 1021 TGTGTCATCTTAGCATGATTTATATGTAATTAATTCGGGACCAACAGGCTTCCTGG 1080
DB 1031 TGTGTCATCTTAGCATGATTTATATGTAATTAATTCGGGACCAACAGGCTTCCTGG 1090
OY 1081 ATCAGCTGTTTATCTCTCTTTTACCAGCAATGATGTTGATCATCGGCTCACAGATA 1140
DB 1091 ATCAGCTGTTTATCTCTCTTTTACCAGCAATGATGTTGATCATCGGCTCACAGATA 1150
OY 1141 TTTTCAAGAGAAATGCGTGGCCGCCACGAGTAAACGTGTCCAGAAATGAAGTTCT 1200
DB 1151 TTTTCAAGAGAAATGCGTGGCCGCCACGAGTAAACGTGTCCAGAAATGAAGTTCT 1210
OY 1201 TACTTACATTAATTTATCAAAATGATGCTGGGTCAAAAGATTTTCTCAAGTGTCA 1260
DB 1211 TACTTACATTAATTTATCAAAATGATGCTGGGTCAAAAGATTTTCTCAAGTGTCA 1270
OY 1261 GAAATATCCGAGAGAGAGAGCTCGATATTTGAAAAAGCGGGGTACTTCCAGAGCATAC 1320
DB 1271 GAAATATCCGAGAGAGAGAGCTCGATATTTGAAAAAGCGGGGTACTTCCAGAGCATAC 1330
OY 1321 TGTGGGTGTGCTCCATTTGTGGTGTATGCCAGCGTGTGACCTTCTGTTCATAT 1380
DB 1331 TGTGGGTGTGCTCCATTTGTGGTGTATGCCAGCGTGTGACCTTCTGTTCATAT 1390
OY 1381 GACCTTGGGCTTCAGTACAGACAGACAGAGCTTTCAGAGGTGACAGCTTCAATTC 1440
DB 1391 GACCTTGGGCTTCAGTACAGACAGACAGAGCTTTCAGAGGTGACAGCTTCAATTC 1450
OY 1441 CATGACTTTTCTTGAAGTAACACCGTTTCACTAAGTCCCTCTCAGAACCTCACT 1500
DB 1451 CATGACTTTTCTTGAAGTAACACCGTTTCACTAAGTCCCTCTCAGAACCTCACT 1510
OY 1501 GGCCTGTTGACAGATTTAAGAGTTTGTCTTAATGGAAGGTTCAATGATTAAGAACAA 1560
DB 1511 GGCCTGTTGACAGATTTAAGAGTTTGTCTTAATGGAAGGTTCAATGATTAAGAACAA 1570
OY 1561 ACCAGCAGATCTCTCAGATCAATCAATAGAGATGAAGAAATGCCACCTTGGGAGCTCTCC 1620
DB 1571 ACCAGCAGATCTCTCAGATCAATCAATAGAGATGAAGAAATGCCACCTTGGGAGCTCTCC 1630
OY 1621 CCACCTCCATATCCAGAAATCTGCCCAAGCTGACCCCAAAATGAAAAAGAACAGAGGC 1680
DB 1631 CCACCTCCATATCCAGAAATCTGCCCAAGCTGACCCCAAAATGAAAAAGAACAGAGGC 1690
OY 1681 TTTCAAGGGGCAAGAAAGAGAGAGTGAAGAGCTGACGCGCACTGAGCATCAGCGGTGCT 1740
DB 1691 TTTCAAGGGGCAAGAAAGAGAGAGTGAAGAGCTGACGCGCACTGAGCATCAGCGGTGCT 1750
OY 1741 GGCAGAGCAGAAAGGCACTCTCTCTGAGACAGTGAAGAGCGGCCAGTCCGAAAGAGA 1800
DB 1751 GGCAGAGCAGAAAGGCACTCTCTCTGAGACAGTGAAGAGCGGCCAGTCCGAAAGAGA 1810
OY 1801 AGAAGGCAAGCAGATCCACCTGGGCGCACTGGGCTTACAGAGAGACACTGCACAGCATGA 1860

```
Db 1811 AGAAGGAGACATCCACCTGGGCGACCTGGCTTACAGAGGACACTGCACACATCGA 1870
Qy 1861 TCTGGATCCAAAGAGGTAATACTGGTATCTGGGCGAGTGTGGGAAGTGGAAAAAC 1920
Db 1871 TCTGGAGATCCAAAGAGGTAATACTGGTATCTGGGCGAGTGTGGGAAGTGGAAAAAC 1930
Qy 1921 CTTCTCATTTACAGCATTTTATAGGCGCAGATGACGCTTCTAGAGGGCAGCATTTGCAATCAG 1980
Db 1931 CTTCTCATTTACAGCATTTTATAGGCGCAGATGACGCTTCTAGAGGGCAGCATTTGCAATCAG 1990
Qy 1981 TGGAACTTGGCTTATATGTGGCGCCAGCAGCGCTGGATCTCTCAATGCTACTCTGAGAGACAA 2040
Db 1991 TGGAACTTGGCTTATATGTGGCGCCAGCAGCGCTGGATCTCTCAATGCTACTCTGAGAGACAA 2050
Qy 2041 CATCTCTTTTGGGAAGGAATATGTAAGAAAGATACAACTCTGTCTGAAACACTGCTG 2100
Db 2051 CATCTCTTTTGGGAAGGAATATGTAAGAAAGATACAACTCTGTCTGAAACACTGCTG 2110
Qy 2101 CTTGAGGCTTGGCTTGGCCATTTCTCCAGCAGCGACCTGACGAGATTTGAGAGCGAGG 2160
Db 2111 CTTGAGGCTTGGCTTGGCCATTTCTCCAGCAGCGACCTGACGAGATTTGAGAGCGAGG 2170
Qy 2161 AGCCAACTTGAGCGGCTGGCGCCAGAGATGAGCTTGGCGGCGCTGTATAGTGA 2220
Db 2171 AGCCAACTTGAGCGGCTGGCGCCAGAGATGAGCTTGGCGGCGCTGTATAGTGA 2230
Qy 2221 CAGGAGCATCTACATCCCTGAGCAGCCCTGAGTCCCTTAGATGCCATGTGGGCAACCA 2280
Db 2231 CAGGAGCATCTACATCCCTGAGCAGCCCTGAGTCCCTTAGATGCCATGTGGGCAACCA 2290
Qy 2281 CATCTTCAATAGTCTTATCCGGAAGCATCTCAAGTCCAGACAGTCTGTGTTTACC 2340
Db 2291 CATCTTCAATAGTCTTATCCGGAAGCATCTCAAGTCCAGACAGTCTGTGTTTACC 2350
Qy 2341 CCATTTACAGTACCTGGTGTGCTGTGATGAAGTATCTTCAAGAGGCGCTATATAC 2400
Db 2351 CCATTTACAGTACCTGGTGTGCTGTGATGAAGTATCTTCAAGAGGCGCTATATAC 2410
Qy 2401 GGAAGAGGCGCCCATGAGAGAACTGATGAATTTAAATGCTATGCTATACATTTTAA 2460
Db 2411 GGAAGAGGCGCCCATGAGAGAACTGATGAATTTAAATGCTATGCTATACATTTTAA 2470
Qy 2461 TAACCTGTGCTGGGAGAGACAGCCGAGTGGATCAATTTCAAAAAAGAAACAGTGG 2520
Db 2471 TAACCTGTGCTGGGAGAGACAGCCGAGTGGATCAATTTCAAAAAAGAAACAGTGG 2530
Qy 2521 TTTACAGAGAAAGTCAACAAGCAAGGGGCTTAAACAGAGATCAATTAAGAGAAAGG 2580
Db 2531 TTTACAGAGAAAGTCAACAAGCAAGGGGCTTAAACAGAGATCAATTAAGAGAAAGG 2590
Qy 2581 AGTAAAGCCAGAGAGAGGCGAGCTTGTGACGCTGGAAGAGAAAGGCGAGGTTACGTC 2640
Db 2591 AGTAAAGCCAGAGAGAGGCGAGCTTGTGACGCTGGAAGAGAAAGGCGAGGTTACGTC 2650
Qy 2641 CTTGCTAGTATATGTGTCTTCAATCCAGGCTGTGCGGGGCGCCCTTGGCATCTGCTTAT 2700
Db 2651 CTTGCTAGTATATGTGTCTTCAATCCAGGCTGTGCGGGGCGCCCTTGGCATCTGCTTAT 2710
Qy 2701 TATGCGCTTTTCAATGCTGATGTAGGAGCAGCGCTTCAACAGCTGTGCTTGAAGTGA 2760
Db 2711 TATGCGCTTTTCAATGCTGATGTAGGAGCAGCGCTTCAACAGCTGTGCTTGAAGTGA 2770
Qy 2761 CTTGATCAAGCAAGGAAAGCGGAAACACCACTGTGATCTGAGGAGCAAGACCTGCTGAG 2820
Db 2771 CTTGATCAAGCAAGGAAAGCGGAAACACCACTGTGATCTGAGGAGCAAGACCTGCTGAG 2830
Qy 2821 TGGACAGATGAAGACATCTCTCATATGCACTACTATGCGAGCATATACGCCCTCTCAT 2880
Db 2831 TGGACAGATGAAGACATCTCTCATATGCACTACTATGCGAGCATATACGCCCTCTCAT 2890
Qy 2881 GGCAGTATGCTGATCTGAAAGGCAATTCGAGGAGTGTCTTTGTCAAAGGCGAGCTGCG 2940
Db 2891 GGCAGTATGCTGATCTGAAAGGCAATTCGAGGAGTGTCTTTGTCAAAGGCGAGCTGCG 2950
Qy 2941 AGCTTCTCCCGGCTGATGACGAGCTTTTCCGAAGATCTTCCGAAGCCCTATGAAGTT 3000
Db 2951 AGCTTCTCCCGGCTGATGACGAGCTTTTCCGAAGATCTTCCGAAGCCCTATGAAGTT 3010
Qy 3001 TTTTGCACGACCCCGACAGGAGGATTTCTACAGGTTTTCCAAAGCATGATGAAGT 3060
Db 3011 TTTTGCACGACCCCGACAGGAGGATTTCTACAGGTTTTCCAAAGCATGATGAAGT 3070
Qy 3061 TGGAGTGGGCTGCGCTTCCAGGCGAGATGTATCATCAGAAAGCTTATCCGCTGCTT 3120
Db 3071 TGGAGTGGGCTGCGCTTCCAGGCGAGATGTATCATCAGAAAGCTTATCCGCTGCTT 3130
Qy 3121 CTTGTGGGAATGATCGAGAGATCTTCCGATGCTCTTGTGGCAGTGGGCGCCCTTGT 3180
Db 3131 CTTGTGGGAATGATCGAGAGATCTTCCGATGCTCTTGTGGCAGTGGGCGCCCTTGT 3190
Qy 3181 CATCTCTTTTCAATCTGTCACATTTGCTCCAGGCTCTGATTTGGGAGCTGAAGCTCT 3240
Db 3191 CATCTCTTTTCAATCTGTCACATTTGCTCCAGGCTCTGATTTGGGAGCTGAAGCTCT 3250
Qy 3241 GGACAAATTCAGCGAGTACACCTTCTCTCCACATCAGTCCAGATTAACGAGGCTTGC 3300
Db 3251 GGACAAATTCAGCGAGTACACCTTCTCTCCACATCAGTCCAGATTAACGAGGCTTGC 3310
Qy 3301 CACCATCCAGCCTACATAAAGGCGAGAGTCTCTGACAGATACGAGAGCTGTGGA 3360
Db 3311 CACCATCCAGCCTACATAAAGGCGAGAGTCTCTGACAGATACGAGAGCTGTGGA 3370
Qy 3361 TGGAAACCAACCTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3420
Db 3371 TGGAAACCAACCTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3430
Qy 3421 GGACCTATCAGATTCGCGCTCATCACACACGCGGCTGATGATCTTCTATGACAGG 3480
Db 3431 GGACCTATCAGATTCGCGCTCATCACACACGCGGCTGATGATCTTCTATGACAGG 3490
Qy 3481 GCAGATTCGCCAGCCTATGCGGCTCTGCGCATCTTATGCTGCAAGTTAACGGGCT 3540
Db 3491 GCAGATTCGCCAGCCTATGCGGCTCTGCGCATCTTATGCTGCAAGTTAACGGGCT 3550
Qy 3541 GTTCCAGTTTACGGTACAGTGGGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
Db 3551 GTTCCAGTTTACGGTACAGTGGGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3610
Qy 3601 GATCAATCACTACATTTAAGACTGTCTTGGAGCACTGCCAGATTTAAGAAAGGCG 3660
Db 3611 GATCAATCACTACATTTAAGACTGTCTTGGAGCACTGCCAGATTTAAGAAAGGCG 3670
Qy 3661 TCCCTCCCTGACTGGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
Db 3671 TCCCTCCCTGACTGGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3730
Qy 3721 CCGAGAAACCTCCCTCCGCTTAAAGAAAGTATCCCTGATCAAACTTAAGAGAA 3780
Db 3731 CCGAGAAACCTCCCTCCGCTTAAAGAAAGTATCCCTGATCAAACTTAAGAGAA 3790
Qy 3781 GATTGGCAATGTGGGCGGAGCAGAGATCAGGAACTCTGCTGGGATGAGCCCTTCCG 3840
Db 3791 GATTGGCAATGTGGGCGGAGCAGAGATCAGGAACTCTGCTGGGATGAGCCCTTCCG 3850
Qy 3841 TCTGTGGAGTATCTGAGGCTGCATCAGATGATGAGTGAATCACTGATATTTGG 3900
Db 3851 TCTGTGGAGTATCTGAGGCTGCATCAGATGATGAGTGAATCACTGATATTTGG 3910
Qy 3901 CTTGCGGAGCTCCGAGCAAACTCTATATCTCTCAAGAGCGGAGCTGCTCAGTGG 3960
Db 3911 CTTGCGGAGCTCCGAGCAAACTCTATATCTCTCAAGAGCGGAGCTGCTCAGTGG 3970
Qy 3961 CACTGTGATCAATTTTGAACCCCTTCAACAGTACAGTGAAGACAGATTTGGAGTGC 4020
Db 3971 CACTGTGATCAATTTTGAACCCCTTCAACAGTACAGTGAAGACAGATTTGGAGTGC 4030
```

OY	244	GTGGAAATGCCAAGATATCCCTTGGAAACACAGCCGAGCCGAGGGCCCTCTCTCTTGATATGC	303
Db	325	GTTCGAATGCCAAGATATCCCTTGGAAACACAGCCGAGCCGAGGGCCCTCTCTTGATATGC	384
OY	304	CTTCATGCATCTTCAGCTCAGAAATCCTGGATAGAGACATCCCAAGGGAAAGTACATCA	363
Db	385	CTTCATGCATCTTCAGCTCAGAAATCCTGGATAGAGACATCCCAAGGGAAAGTACATCA	444
OY	364	TGGCTTGAAGTCCTTGAAGCCCATCCGGACTACCTTGCACAAACACAGCACCAAGTGGACAA	423
Db	445	TGGCTTGAAGTCCTTGAAGCCCATCCGGACTACCTTGCACAAACACAGCACCAAGTGGACAA	504
OY	424	TGCTGGGCTTTTTCCTGATATGACTTTTTCGTGAGCTTCTTCTTGCGCCCTGTGGCCCA	483
Db	505	TGCTGGGCTTTTTCCTGATATGACTTTTTCGTGAGCTTCTTCTTGCGCCCTGTGGCCCA	564
OY	484	CAAGAAGGGGAGGAGCTCATATGGAAGACGATGATCTGTGCACAGACAGAGTCCTTGCGA	543
Db	565	CAAGAAGGGGAGGAGCTCATATGGAAGACGATGATCTGTGCACAGACAGAGTCCTTGCGA	624
OY	544	CGTGAACATGCAGAAAGACTATAGAGACTGTGGCAAGAAGAGCTGAATGAAGTTGGGCGAGA	603
Db	625	CGTGAACATGCAGAAAGACTATAGAGACTGTGGCAAGAAGAGCTGAATGAAGTTGGGCGAGA	684
OY	604	CGCTGCTTCCTCTGGCAAGGGGTGTGTGATCTTTCGCGCACAGGCTCATCTGTCCAT	663
Db	685	CGCTGCTTCCTCTGGCAAGGGGTGTGTGATCTTTCGCGCACAGGCTCATCTGTCCAT	744
OY	664	CGTGTGCTGATATATCACAGCTGGCTTGAGTTCAGTGGACACAGCCCTTCATGTGTAACA	723
Db	745	CGTGTGCTGATATATCACAGCTGGCTTGAGTTCAGTGGACACAGCCCTTCATGTGTAACA	804
OY	724	CCCTTGGAGATATCCCAAGCAACAGATCTTAACCTGCAGTACACTTGTGTTAGTACT	783
Db	805	CCTCTTGGAGATATCCCAAGCAACAGATCTTAACCTGCAGTACACTTGTGTTAGTACT	864
OY	784	GGGCGCTCCTCTGACGGAAATCGTGGCGTCTTGCTGCTTGCACTGACTTGGGCAATTGAA	843
Db	865	GGGCGCTCCTCTGACGGAAATCGTGGCGTCTTGCTGCTTGCACTGACTTGGGCAATTGAA	924
OY	844	TTTACCGAAACCGGTGTCGGCTTGGGGGGGCGCATCTTAACCATATGGCATTTTAAAGATATCT	903
Db	925	TTTACCGAAACCGGTGTCGGCTTGGGGGGGCGCATCTTAACCATATGGCATTTTAAAGATATCT	984
OY	904	TAAGTTAAAGAACTTTAAAGAAATTCCTGGGGTGAGCTATCAACAATTTGCTCCACAGA	963
Db	985	TAAGTTAAAGAACTTTAAAGAAATTCCTGGGGTGAGCTATCAACAATTTGCTCCACAGA	1044
OY	964	TGGCGAGAGAAATGTTTGAAGCAGACAGCCGTGTGGACACCTGTGCGGTGGAGACCCGTTGT	1023
Db	1045	TGGCGAGAGAAATGTTTGAAGCAGACAGCCGTGTGGACACCTGTGCGGTGGAGACCCGTTGT	1104
OY	1024	TGCCATCTTGAAGCATGATTTTAATATGTAATTAATCTGGGACCAACAGGCTTCTTGCGATC	1083
Db	1105	TGCCATCTTGAAGCATGATTTTAATATGTAATTAATCTGGGACCAACAGGCTTCTTGCGATC	1164
OY	1084	AGCTGTTTTTATCCCTTTTAAACCAGCAATGATGTTTGATATACAGGCTCACACGATATTT	1143
Db	1165	AGCTGTTTTTATCCCTTTTAAACCAGCAATGATGTTTGATATACAGGCTCACACGATATTT	1224
OY	1144	CAGGAGAAATGCTGTGGCCGCCACAGGTAGCTGTGCACAGAGATGAATGAAGTCTTAC	1203
Db	1225	CAGGAGAAATGCTGTGGCCGCCACAGGTAGCTGTGCACAGAGATGAATGAAGTCTTAC	1284
OY	1204	TTTACATTTAAATTTATCAAAATGTAATGCTGTGGTCAAAAGCATTTTCTCAGAGTGTTCAGAA	1263
Db	1285	TTTACATTTAAATTTATCAAAATGTAATGCTGTGGTCAAAAGCATTTTCTCAGAGTGTTCAGAA	1344
OY	1264	AATCCGCGAGAGAGAGGTGCGAATTTTGGAAAAAGCCGGGTACTTCCAGAGCATACAGT	1323
Db	1345	AATCCGCGAGAGAGAGGTGCGAATTTTGGAAAAAGCCGGGTACTTCCAGAGCATACAGT	1404
OY	1324	GGGTGTGGCTCCCAATTTGTGTGTGATTTGGCACGCTGTGATCACTTCTCTGTTCAATATAC	1383

Db	1405	GGGTGTGGCTCCATTGTGGTGTATTTGGCAGCGGTGGACCTTCTGTTCATATGAC	1464
Qy	1384	CTGTGGCTTCGATCTGCACAGCACAGAGCTTTCACAGTGTGACAGTCTTCAATTCCAT	1443
Db	1485	CTGTGGCTTCGATCTGCACAGCACAGAGCTTTCACAGTGTGACAGTCTTCAATTCCAT	1524
Qy	1444	GACTTTCCTTTCGAAAGTAACACCGTTTTCAGTAAAGTCCCTCTCAGAAAGCTCAGTGGC	1503
Db	1525	GACTTTCCTTTCGAAAGTAACACCGTTTTCAGTAAAGTCCCTCTCAGAAAGCTCAGTGGC	1584
Qy	1504	TGTTGACAGATTTAAGGTTTGTTCCTAATGGAAGAGGTTCACATGATAAAGAACAAAC	1563
Db	1585	TGTTGACAGATTTAAGGTTTGTTCCTAATGGAAGAGGTTCACATGATAAAGAACAAAC	1644
Qy	1564	AGCCAGTCTCATCATCAAGATTAAGATGAAAAAATGGCACCTTGGCATGGAGCTCTCCCA	1623
Db	1645	AGCCAGTCTCATCATCAAGATTAAGATGAAAAAATGGCACCTTGGCATGGAGCTCTCCCA	1704
Qy	1624	CTCCAGTATCCAGAACTCGCCCAAGCTGACCCCCCAAAATGAAAAAAGCACAGAGGGCTTC	1683
Db	1705	CTCCAGTATCCAGAACTCGCCCAAGCTGACCCCCCAAAATGAAAAAAGCACAGAGGGCTTC	1764
Qy	1684	CAGGGCAGAAAGAGAGAGTGAAGCAGCTGCACAGCGCATGAGCATCAGCGGGTCTGGC	1743
Db	1765	CAGGGCAGAAAGAGAGAGTGAAGCAGCTGCACAGCGCATGAGCATCAGCGGGTCTGGC	1824
Qy	1744	AGAGCAGAAAGCCACCTCCCTCCGTCGACAGTGCAGAGGGGGCCACATCCCGAAGAGAGA	1803
Db	1825	AGAGCAGAAAGCCACCTCCCTCCGTCGACAGTGCAGAGGGGGCCACATCCCGAAGAGAGA	1884
Qy	1804	AGGCAGACATCTCCACCTGGGGCCACCTGGCGCTTACAGAGACACTGCACAGATCGATCT	1863
Db	1885	AGGCAGACATCTCCACCTGGGGCCACCTGGCGCTTACAGAGACACTGCACAGATCGATCT	1944
Qy	1864	GGAGATCCACAGAGGTAACCTGTTGGATCTTGGCAGTGTGGAGTGGAAAACTC	1923
Db	1945	GGAGATCCACAGAGGTAACCTGTTGGATCTTGGCAGTGTGGAGTGGAAAACTC	2004
Qy	1924	TCTCATTTCAACCAATTTAGGCCAGATGACGCTTTCAGAGGGCCACATTTCAATAGTGG	1983
Db	2005	TCTCATTTCAACCAATTTAGGCCAGATGACGCTTTCAGAGGGCCACATTTCAATAGTGG	2064
Qy	1984	AACCTTCGCTTATGTGGCCAGCAGAGGCTTGATCTCAATGCTACTGTAGAGACATAT	2043
Db	2065	AACCTTCGCTTATGTGGCCAGCAGAGGCTTGATCTCAATGCTACTGTAGAGACATAT	2124
Qy	2044	CTGTGTTGGGAAGAAATATGATGAAGAAATCAACTGTGTGTCGAACAGCTGTGCT	2103
Db	2125	CTGTGTTGGGAAGAAATATGATGAAGAAATCAACTGTGTGTCGAACAGCTGTGCT	2184
Qy	2104	GAGGCTTCACTGGCCCAATCTTCTCCACAGCACCTGCAGAGGAGATTTGGAGAGCAGAGC	2183
Db	2185	GAGGCTTCACTGGCCCAATCTTCTCCACAGCACCTGCAGAGGAGATTTGGAGAGCAGAGC	2244
Qy	2164	CAACCTTGAGCGGTGGGACAGGCCAGAGAGATCACCTTGGCCGGGCTTGTATAGTACAG	2223
Db	2245	CAACCTTGAGCGGTGGGACAGGCCAGAGAGATCACCTTGGCCGGGCTTGTATAGTACAG	2304
Qy	2224	GAGCATCTACATCTCTGGACAGACCCCTCACTGCTTGAATGCCATGTGGGCAACACAT	2283
Db	2305	GAGCATCTACATCTCTGGACAGACCCCTCACTGCTTGAATGCCATGTGGGCAACACAT	2364
Qy	2284	CTTCAATAGTCTATCCGGGAAACATCTCAAGTCCAGAGACAGATTTCTTTTGTATCCACCA	2343
Db	2365	CTTCAATAGTCTATCCGGGAAACATCTCAAGTCCAGAGACAGATTTCTTTTGTATCCACCA	2424
Qy	2344	GTTACAGTACCTGGTTTTCACGTGATGAAGTCAATCTCATGAAGAAGGCGTGTATTACGA	2403
Db	2425	GTTACAGTACCTGGTTTTCACGTGATGAAGTCAATCTCATGAAGAAGGCGTGTATTACGA	2484
Qy	2404	AAGAGCACCCTATGAGACTGATGAATTTAAATGGTACTATGCTACCATTTTAAATA	2463

QY 4021 CCTGAGAGGACACATGAAAGATGATTCCTAGCTACCTCTGAACCTGAATCTGA 4080
 DB 4031 CCTGAGAGGACACACATGAAAGATGATTCCTAGCTACCTCTGAACCTGAATCTGA 4090
 QY 4081 AGTGAATGAGAAATGGGGAATACCTCTAGTGGGGAACGGAGCTCTTGATGATAGTAG 4140
 DB 4091 AGTGAATGAGAAATGGGGAATACCTCTAGTGGGGAACGGAGCTCTTGATGATAGTAG 4150
 QY 4141 AGCCCTGCTCCGCCACTGTAAGATTCGATTTTAATGAGGACAGAGCTGCCATGGACAC 4200
 DB 4151 AGCCCTGCTCCGCCACTGTAAGATTCGATTTTAATGAGGACAGAGCTGCCATGGACAC 4210
 QY 4201 AGAGACAGACTTATGATTCAGAGACATCCGAGAGACTTGGAGACTGATACATGCT 4260
 DB 4211 AGAGACAGACTTATGATTCAGAGACATCCGAGAGACTTGGAGACTGATACATGCT 4270
 QY 4261 GACCAATGCCCATTCCTGTCAGAGAACAGAGCTGTCAGAGGCTGACTCCCTGCTGT 4320
 DB 4271 GACCAATGCCCATTCCTGTCAGAGAACAGAGCTGTCAGAGGCTGACTCCCTGCTGT 4330
 QY 4321 GGGAGAGGTGGTGGAGTTGACACCCCATCGCTCTGTCGCAACGAGATGCCGAT 4380
 DB 4331 GGGAGAGGTGGTGGAGTTGACACCCCATCGCTCTGTCGCAACGAGATGCCGAT 4390
 QY 4381 CATGCCATGTTGCTGTCAGAGAACAGAGCTGTCAGAGGCTGACTCCCTGCTGT 4440
 DB 4391 CATGCCATGTTGCTGTCAGAGAACAGAGCTGTCAGAGGCTGACTCCCTGCTGT 4450
 QY 4441 TGACGAGTCTCTTCTTTAGAGCATTCGATCCCTGCTGCGGGGGGGCCCTTCATC 4500
 DB 4451 TGACGAGTCTCTTCTTTAGAGCATTCGATCCCTGCTGCGGGGGGGCCCTTCATC 4509
 QY 4501 GCGTCTCTTACCGAACTGCTTCTGCAATTTTTCGCAACGAGATGCCGAT 4560
 DB 4510 GCGTCTCTTACCGAACTGCTTCTGCAATTTTTCGCAACGAGATGCCGAT 4569
 QY 4561 TGGCTGTGCTGCTTACCTTTAGGAGAGTCAATTTTGTATTTATTTATTCATAT 4620
 DB 4570 TGGCTGTGCTGCTTACCTTTAGGAGAGTCAATTTTGTATTTATTTATTCATAT 4629
 QY 4621 TCATGTAAACAAATTTAGTTTGTCTTAAATGCACTCTTAAAGAGTTCAGGAAACCT 4680
 DB 4630 TCATGTAAACAAATTTAGTTTGTCTTAAATGCACTCTTAAAGAGTTCAGGAAACCT 4689
 QY 4681 TATTTATTTATTTAGAGAGGCTTATTAATGAGAGCTTATTAATGAGAGCTTATTAAT 4740
 DB 4690 TATTTATTTATTTAGAGAGGCTTATTAATGAGAGCTTATTAATGAGAGCTTATTAAT 4749
 QY 4741 AATCTGTACATAGCTTATTTATTAAGTGAATAATGTAAGCTTTATTTATTTATTAAT 4800
 DB 4750 AATCTGTACATAGCTTATTTATTAAGTGAATAATGTAAGCTTTATTTATTTATTAAT 4809
 QY 4801 AAGCACTGTCTTAAAAAA 4819
 DB 4810 AAGCACTGTCTTAAATACA 4828

RESULT 3
 LOCUS AF146074 5881 bp mRNA PRI 10-AUG-1999
 DEFINITION Homo sapiens ABC protein mRNA, complete cds.
 ACCESSION AF146074
 VERSION AF146074.1 GI:5006890
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 5881)
 McAlister, M.A., Breen, M.A., White, N.L. and Matthews, N.
 PAC11 (also known as MOAT-C and MRP5), a member of the ABC family
 of proteins, has anion transporter activity but does not confer

multiliding resistance when overexpressed in human embryonic kidney
 293 cells
 J. Biol. Chem. 274 (33), 23541-23548 (1999)
 MEDLINE 99367488
 REFERENCE 2 (bases 1 to 5881)
 McAlister, M.A., Breen, M., White, N. and Matthews, N.
 AUTHORS Direct Submission
 TITLE Submitted (27-Apr-1999) Yamanouchi Research Institute, Armstrong
 JOURNAL Road, Oxford OX4 45X, UK
 FEATURES Location/Qualifiers
 source 1..5881
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /tissue_type="Brain"
 /dev_stage="adult"
 /protein_id="AAC37716.1"
 /product="ABC protein"
 /note="PAC11; MOAT-C; MRP5"
 /codon_start=1
 /function="anion transporter"
 /note="PAC11; MOAT-C; MRP5"
 /codon_start=1
 /product="ABC protein"
 /protein_id="AAC37716.1"
 /db_xref="GI:5006891"
 /translation="MKDIDIGKEYIIPSGYRSYREKTSIGTHREDKSRRTPL
 ECODALEPAAAREGLSDASMSQRLIDEPHPGKHGSLAKPIRTTSHOHVVD
 NAGLESCMTFEMLSLARVARKKGLSMEVWSLSKHSVNCRLERLQOELENEY
 GPDAISLRVWICRRLILSIYCIIMTOLAGSGPFWYKHLLETLOQAEINLOYS
 LLIVGLITPEIVSMALITWALNPRGVRLKRAITLMARKITIKLNKIKESIGEL
 INISNDGQRFEEAAGVSLAGGEVAILLHMTINVLITGTEGLGSAVFLIETPAM
 FASRLTAYERRKCAVADEROKMNEVITYIKFIMAWVAEFSQSVKFIETRRIL
 EKAGTFOSTIVGVAPIVVIVIASVYFVHMGLDLPAAQFTVYVENSMTLAKVT
 PFSVYSLSIEAVADRFEKSLFIMEVHMKPKSPHKKIEMKATLAMDSSHSION
 SPKLTPMKMKRKRASRGKKEKROLOREHOVALEOKGHLILDSDESPDEEGKH
 IHLRLIORTLHVSDELIOBKILVIGCGSGKTSLSIILDOMLRETSISGT
 FAYVAQAMILNATLRDNLIPKEXDERNSVYNSCLRPDLAILSSDTEIGERG
 ANLSGGQQRSLNLAALYSDSRYIILDPISLALAHVGNHIFNSAIRKHLKSVLFV
 THOLOYLVDCEVIEFMKEDCTIERGTHELMNLNGDAVITFNMLIGETPEVLENSK
 ETSGSKKSOQKGRKTSVKKKAVKPEEGLOVLEEGQGSVMSYGVYIOAAGP
 LAPVIMALFPLNLTNGSTAFSTFWMLSYWKOGSGNTYVRGNETSVDSMKPMHOY
 ASIALSAMVILIKATIRGVFVCTLPASRLDELPRRLRBRMPEPTTPGRLI
 NRSKDMDEVRLDPQAEMLIOWIIVFVCGMAGTFPFVAVGPEVILVLSVLI
 VSRVILRELRLNDTOSPLSHYTSISIOGLATIHAYKGOEFLHROELLDNDQAP
 FLFTICAMRLAVRLDISIALITTTGLMIVHMOIIPAYAGLISAVOLJTGFOFT
 VRLASEPRLSVRIINHYIKTSLLEPAIKKAPSDPMDPGEVFEENAEKRYE
 NLPLIKKVSPTIKPKKIGIVGRGSGSLGMAFLPEVLSGGCIKIDGVRISDGL
 LADLRKSLIIPBQVLESGTVRSNLDPEFNOYTEDOIMDALERTHMKECIADLPKLE
 SEVMENGNESVGEROLICTARALLRHKILILDEBATRAMDETDLIOETIRPAD
 CTMLTIAHRLHTVLGSDRIMVLAOGQVEFDPVSLNDSRFYAMFAAENVAAYK
 G"

BASE COUNT 1414 a 1448 c 1531 g 1488 t
 ORIGIN
 Query Match 98.8%; Score 4789.6; DB 89; Length 5881;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 4806; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
 QY 4 TCATGTCGGAGCGGTGTGAGCGCGGCTGCTGCTGAGAGGCGGAGGAA 63
 DB 85 TAAATCTCGGAGCGGTGTGAGCGCGGCTGCTGCTGAGAGGCGGAGGAA 144
 QY 64 TTCTGATGGAAGTAAACAGTCTGTGAGCGCTGGAACCTCCACTAGAGAAATGAAGA 123
 DB 145 TTCTGATGGAAGTAAACAGTCTGTGAGCGCTGGAACCTCCACTAGAGAAATGAAGA 204
 QY 124 TATCGACATAGGAAAAAGACTATATCATCCCAAGTCTGTGAGTAAAGTGTGAGGAGAG 183
 DB 205 TATCGACATAGGAAAAAGACTATATCATCCCAAGTCTGTGAGTAAAGTGTGAGGAGAG 264
 QY 184 AACCAAGACTTTCTGGAGCGGACAGAGACCGGTGAAGATTCAGAGTTCAGAGAACTCGACC 243
 DB 265 AACCAAGACTTTCTGGAGCGGACAGAGACCGGTGAAGATTCAGAGTTCAGAGAACTCGACC 324

QY 4624 TGTAAACAAATTTAGTTTGTCTTAATTCACCTCTAAAGGTTGAGGACCGTTAT 4683
|||||
Db 4704 TGTAAACAAATTTAGTTTGTCTTAATTCACCTCTAAAGGTTGAGGACCGTTAT 4763
QY 4684 TATTAATGTATCAGAGCCCTTAATGAAGCTTTATACCTAGCTATATATATAT 4743
|||||
Db 4764 TATTAATGTATCAGAGCCCTTAATGAAGCTTTATACCTAGCTATATATATAT 4823
QY 4744 TCTGTACATACCTATATTTTACATGAAATGATGAGCTTTATTTATATTAATAG 4803
|||||
Db 4824 TCTGTACATACCTATATTTTACATGAAATGATGAGCTTTATTTATATTAATAG 4883
QY 4804 CACTGTGCTATAAAAAA 4819
|||||
Db 4884 CACTGTGCTATAAACA 4899
RESULT 4
HSU83661 5826 bp mRNA PRI 21-JUN-2000
LOCUS Homo sapiens multidrug resistance protein 5 (MRP5) mRNA, complete
DEFINITION cds.
ACCESSION U83661
VERSION U83661.2 GI:5685863
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 3765 to 5822)
AUTHORS Kool,M., de Haas,M., Scheffer,G.L., Scheper,R.J., van Eijk,M.J.,
Juijn,J.A., Baas,F., and Borst,P.
TITLE Analysis of expression of cMOAT (MRP2), MRP3, MRP4, and MRP5,
homologues of the multidrug resistance-associated protein gene
(MRP1), in human cancer cell lines
JOURNAL Cancer Res. 57 (16), 3537-3547 (1997)
MEDLINE 97413640
REFERENCE 2 (bases 1 to 5826)
AUTHORS Wijnholds,J., Mol,C.A., van Deemter,L., de Haas,M., Scheffer,G.L.,
Baas,F., Beijnen,J.H., Scheper,R.J., Hulse,S., De Clercq,E.,
Balzarini,J., and Borst,P.
TITLE Multidrug-resistance protein 5 is a multispecific organic anion
transporter able to transport nucleotide analogs
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (13), 7476-7481 (2000)
MEDLINE 20319051
REFERENCE 3 (bases 3765 to 5822)
AUTHORS Kool,M., Wijnholds,J., de Haas,M., Mol,C.A.A.M., Ponne,N.J.,
Baas,F., and Borst,P.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-1997) Molecular Biology, Netherlands Cancer
Institute, Plesmanlaan 121, Amsterdam 1066 CX, The Netherlands
REFERENCE 4 (bases 1 to 5826)
AUTHORS Wijnholds,J., Mol,C.A.A.M., Ponne,N.J., Baas,F., and Borst,P.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Molecular Biology, Netherlands Cancer
Institute, Plesmanlaan 121, Amsterdam 1066 CX, The Netherlands
REMARK Sequence update by submitter
COMMENT On Aug 3, 1999 this sequence version replaced gi:2439973.
FEATURES
source
1..5826
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="7.69 cR from WI-6365"
/tissue_type="brain"
/dev_stage="fetus"
1..5826
/gene="MRP5"
120..4433
/gene="MRP5"
/note="ABC transporter protein; MOAT-C; PABCI1"
/codon_start=1
/product="multidrug resistance protein 5"

/protein_id="AA871758.2"
/db_xref="GI:5685864"
/translation="MKDIDIGKEVYIIIPSGYRSVREKRTSNQGRHREDSKERRRPL
ECDDALETARABGLSDISMHSQILHIDBEHKGRKHHLSALKRIITTSKIQHYD
NAGLFSCMTFSWLISLARVAHKRGELISMEDVWSLSHESDVNKLREYTOATELQDS
GPDASLRVWVWIECRRLILSVCLMTLOAFSGPAFWVKHLLETOATELQDS
LLVLVLGILLLEIVRSWLSALTALNRTGRLGALITLMAFKILILKLNKESLDEL
INICSDNGOMPEFAAAGVSLLAGPVVALIGMILYNYILIGPTGFLSAVILFYPM
FASRLVAFFRRKCVAAVDERVKMNEVLYIRIKIYAAWKAQSQSVKIFREERL
EKAGYQSTIVYGAIPYVYIAYVYTSVYHNTLFDITAAQAPVYVYVNSMTALVY
PFVSKLSEASVAVDRKSLFLMEVHMINKRASPITKEMNATLWDSSTION
SPRLTPMKKDKRASRKKERKQVLOQRTHEQAVLAQKGLLDSDERSPEDESKH
ILHGLRLQRTLSIDLEIOEGKLVGSGSVGSKLSLSAILIGOMTLLENGSLAIGT
FAYVAQAOAILNATLNDNLNILEGKEYBERNSVNSCCLPDAIPLSSDLTGERG
ANLSGGOORILSARALYSRSLYIIDDPISALDAHNGHIFNSAIRKHLKSTVLVY
THOLOLVDDDEYIPKKEGTYIRGRHLEELMANGXYATFPNNLLGEPFVINSK
ETISGSKSODKGPKTGSYKKAVERKBEQGLVLEKRGGSVSWYGYIYDAAQGP
LAPLVIALFMLELVGSAFSTWMLSWIKGSGNTVYRQENETVSQSMKDNHMOY
ASIVALSMAVMLILKAIKGVVFKGTLRASSRLHDELFRILRSPMKFPDPTPTRL
NRESKMDDEVADRLPQAEFIONVLIVFSCVGMILAGVFWFLVAVGLVILFSLVH
VSRVILIRELRLDNLITQSPFLSHSTTSIOGLATIHAYNKGOELHRYOELDNOAF
FLTPCARMVLAFLDLISALITTTGMLTYLMGQIPRAYAGLISVAYOGLRFOY
VRLASEEARFTVERLNHITKLSLEAPRIKNAKSPDMPDGEVTFENAKRIE
MLPLVLRKVSFTIKPKKIGIVRGTSKSLGMALFRLVELGGCITKIDGVRISDIG
LADLRSKLSIIPQEPVLFSGTVRSNLDPEQYEDQIMDALERTHEKECIAQPLKLE
SEVMDNFSVEBROLICARALLRICKILIDEATAAMDTELDLLOETIRLEAFAD
CTMLTIAHRLHYVIGSDRIWVLAQGOVYEDPFSVLSNDSRFYAMFAAEKNAVYK
G"

BASE COUNT 1404 a 1427 c 1511 g 1474 t 10 others
ORIGIN

Query Match 98.8%; Score 4788; DB 97; Length 5826;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4802; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 12 GGGAGCGTGTTGAGCGCGCTGCCGCGTTCCTGAGACAGGCGCAGAAATTCGTATG 71
|||||
Db 16 GGGAGCGTGTTGAGCGCGCTGCCGCGTTCCTGAGACAGGCGCAGAAATTCGTATG 75
QY 72 TGAACATACAGTCTGTGAGCCCTGGAACTCCACTCAGAGAGATGAGATATCGACA 131
|||||
Db 76 TGAACATACAGTCTGTGAGCCCTGGAACTCCACTCAGAGAGATGAGATATCGACA 135
QY 132 TAGGAAAAGATATATATATCCCACTCGTGGGTATAGAAAGTGAAGGAGAGAACCGACA 191
|||||
Db 136 TAGGAAAAGATATATATATCCCACTCGTGGGTATAGAAAGTGAAGGAGAGAACCGACA 195
QY 192 CTTCTGGAGCGCACAGAGACCGTGAAGATTCAGATTCAGAGAACTCGACCGTTGGAT 251
|||||
Db 196 CTTCTGGAGCGCACAGAGACCGTGAAGATTCAGATTCAGAGAACTCGACCGTTGGAT 255
QY 252 GCCAAGATGCTTGGAAAACAGACGCCGAGGCGCTCTCTTTGATGCTTCATATG 311
|||||
Db 256 GCCAAGATGCTTGGAAAACAGACGCCGAGGCGCTCTCTTTGATGCTTCATATG 315
QY 312 ATTCTCAGCTAGAAATCTCGATGAGAGACATCCCAAGGAAATAGCATATGCTTGA 371
|||||
Db 316 ATTCTCAGCTAGAAATCTCGATGAGAGACATCCCAAGGAAATAGCATATGCTTGA 375
QY 372 GTGCTGTGAAGCCCATCGGACTACTTGGCAAAACACAGACCCAGTGAACATCTCGGC 431
|||||
Db 376 GTGCTGTGAAGCCCATCGGACTACTTGGCAAAACACAGACCCAGTGAACATCTCGGC 435
QY 432 TTTTTCCTGTATGACTTTTCTGTGGCTTCTCTGCGCCGTGTGGCCACAGAAAG 491
|||||
Db 436 TTTTTCCTGTATGACTTTTCTGTGGCTTCTCTGCGCCGTGTGGCCACAGAAAG 495
QY 492 GGGAGCTCTCAATGAGAGAGCTGTGCTCTGTCAGCAAGACATCTTCAGCTGAAGT 551
|||||
Db 496 GGGAGCTCTCAATGAGAGAGCTGTGCTCTGTCAGCAAGACATCTTCAGCTGAAGT 555
QY 552 GCAGAGACTGTAGAGAGACTGTGCGAAGAGCTGAATGAAGTTGGGCGCAGACGCTGT 611

```
Db 2485 AAGAGGCAACCCATGAGAGCACTGATTAATTTAAATGCTACTATGCTACCATTTTAAATA 2544
QY 2464 CCTGTTGCTGGAGAGACACCGCCAGTTGAGATCAATTCAAAAAGAAACCAAGTGGTTC 2523
Db 2545 CCTGTTGCTGGAGAGACACCGCCAGTTGAGATCAATTCAAAAAGAAACCAAGTGGTTC 2604
QY 2534 ACAGAGAAGTCAACAGACACAGGGTCTTAAACAGAGATCAATTAAGAGAAAGAAAGCACT 2583
Db 2605 ACAGAGAAGTCAACAGACACAGGGTCTTAAACAGAGATCAATTAAGAGAAAGAAAGCACT 2664
QY 2584 AAAGCCAGAGAGAGGGAGCTGTGACGTGGAGAGAAAGGGGAGGGTTCAGTGGCCCG 2643
Db 2665 AAAGCCAGAGAGAGGGAGCTGTGACGTGGAGAGAAAGGGGAGGGTTCAGTGGCCCG 2724
QY 2644 GTCACTATATGCTGTCTACATCCAGGCTGCTGGGGCCCTTGCGCATTCCTGTTATAT 2703
Db 2725 GTCACTATATGCTGTCTACATCCAGGCTGTGGGGCCCTTGCGCATTCCTGTTATAT 2784
QY 2704 GGCCCTTTTCATGCTGAATGTAGGACACCGCCTTCAGCACCTGGTGGTGAATTA 2763
Db 2785 GGCCCTTTTCATGCTGAATGTAGGACACCGCCTTCAGCACCTGGTGGTGAATTA 2844
QY 2764 GATCAAGCAAGAGAAACGGGAAACACCACTGTGACTGAGAGGAAACAGACCTCGGTAGTGA 2823
Db 2845 GATCAAGCAAGAGAAACGGGAAACACCACTGTGACTGAGAGGAAACAGACCTCGGTAGTGA 2904
QY 2824 CAGCATGAAGAGCAATCCATCATATGACATGACATGACATCATGACCTCTCCATGAC 2883
Db 2905 CAGCATGAAGAGCAATCCATCATATGACATGACATGACATCATGACCTCTCCATGAC 2964
QY 2884 AGTCATGCTGATCTGAAAGCCATTCGAGAGATGTCTTTGTCAAGGGGACGCTGGAGC 2943
Db 2965 AGTCATGCTGATCTGAAAGCCATTCGAGAGATGTCTTTGTCAAGGGGACGCTGGAGC 3024
QY 2944 TTTCCTCCCGGCTGCTGATGACGACCTTTCCGAGAGATCCTTCGAGAGCCCTATGAATTT 3003
Db 3025 TTTCCTCCCGGCTGCTGATGACGACCTTTCCGAGAGATCCTTCGAGAGCCCTATGAATTT 3084
QY 3004 TGACAGACACCCCAAGAGAGATCTCAACAGTTTCCAAAGACATGATGAATGA 3063
Db 3085 TGACAGACACCCCAAGAGAGATCTCAACAGTTTCCAAAGACATGATGAATGA 3144
QY 3064 CGTGGGGCTGCCGTTCCAGGCCGAGATGTTCAATCCAGAAAGCTTATCCTGTTCTG 3123
Db 3145 CGTGGGGCTGCCGTTCCAGGCCGAGATGTTCAATCCAGAAAGCTTATCCTGTTCTG 3204
QY 3124 TTGGGGAATGATCGAGAGAGTCTCCCGGTGCTCTGTGGGAGAGGGCCCTGTGAT 3183
Db 3205 TTGGGGAATGATCGAGAGAGTCTCCCGGTGCTCTGTGGGAGAGGGCCCTGTGAT 3264
QY 3184 CCTTTTTCAGTCTGACATGTTCTCCAGGTCCGATTTCCGGAGCTGAACGCTGGA 3243
Db 3265 CCTTTTTCAGTCTGACATGTTCTCCAGGTCCGATTTCCGGAGCTGAACGCTGGA 3324
QY 3244 CAATATACAGCATCCTTTCTCTCCACATCAAGTCCAGCATACAGGCTTGGCAC 3303
Db 3325 CAATATACAGCATCCTTTCTCTCCACATCAAGTCCAGCATACAGGCTTGGCAC 3384
QY 3304 CATTCACGCTCAATTAAGGGGAGAGTTCGACAGATACAGAGACCTCTGATGA 3363
Db 3385 CATTCACGCTCAATTAAGGGGAGAGTTCGACAGATACAGAGACCTCTGATGA 3444
QY 3364 CAACCAAGCTCTTTTCTTTTGTTAAGTGTGAGTGGGTGGGTGGGTGGGCTGGA 3423
Db 3445 CAACCAAGCTCTTTTCTTTTGTTAAGTGTGAGTGGGTGGGTGGGTGGGCTGGA 3504
QY 3424 CCTTCATCAGCATCGCCCTTCATCACAACAGGGGCTGATGATGTTCTTATCAGAGGGA 3483
Db 3505 CCTTCATCAGCATCGCCCTTCATCACAACAGGGGCTGATGATGTTCTTATCAGAGGGA 3564
QY 3484 GATTCGCCAGGCTATGGGGGTCTGGCCATCTCTTATGCTGTCCAGTTAACGGGGCTGTT 3543
Db 3565 GATTCGCCAGGCTATGGGGGTCTGGCCATCTCTTATGCTGTCCAGTTAACGGGGCTGTT 3624

QY 3544 CCAATTTACGTTACAGCTGGCATCTGAGACAGAGCTGATTCACCTGGTGGAGAGAT 3603
Db 3625 CCAATTTACGTTACAGCTGGCATCTGAGACAGAGCTGATTCACCTGGTGGAGAGAT 3684
QY 3604 CAATCACTACATTAAGACTCTGCTGGAAGACACTGTCACAAATTAAGACAGGCTTC 3663
Db 3685 CAATCACTACATTAAGACTCTGCTGGAAGACACTGTCACAAATTAAGACAGGCTTC 3744
QY 3664 CTCCTCTGACTGCCCCAGAGAGAGAGTGAACCTTTGAAACGCAAGATGAGTACCG 3723
Db 3745 CTCCTCTGACTGCCCCAGAGAGAGAGTGAACCTTTGAAACGCAAGATGAGTACCG 3804
QY 3724 AGAAAACTTCCTCTCTGCTTAAAGAAATATCTTCACGATCAAACTTAAGAGAAAT 3783
Db 3805 AGAAAACTTCCTCTCTGCTTAAAGAAATATCTTCACGATCAAACTTAAGAGAAAT 3864
QY 3784 TGGCATTTGGGGCGGACAGAGATCAGGGAAATCTGCTGGGGATGGCCCTCTCCGTCT 3843
Db 3865 TGGCATTTGGGGCGGACAGAGATCAGGGAAATCTGCTGGGGATGGCCCTCTCCGTCT 3924
QY 3844 GGTGAGATTATCTGGAGGCTGCAATCAAGATGATGAGTGAATCAGTATATTGGCCT 3903
Db 3925 GGTGAGATTATCTGGAGGCTGCAATCAAGATGATGAGTGAATCAGTATATTGGCCT 3984
QY 3904 TGCCGACCTCCGAAGCAAACTCTATTCATCTCTCAAGAGCCGGTCTGTCAGTGGCAC 3963
Db 3985 TGCCGACCTCCGAAGCAAACTCTATTCATCTCTCAAGAGCCGGTCTGTCAGTGGCAC 4044
QY 3964 TGTCAATCAAAATTTGGAGCCCTTCAACCAATGACTGAAGACAGATTTGGGATGCCCT 4023
Db 4045 TGTCAATCAAAATTTGGAGCCCTTCAACCAATGACTGAAGACAGATTTGGGATGCCCT 4104
QY 4024 GGAGAGACACACATGAAGAAATGTATGCTCAGCTACCTCGAAACTGAATCTGAAT 4083
Db 4105 GGAGAGACACACATGAAGAAATGTATGCTCAGCTACCTCGAAACTGAATCTGAAT 4164
QY 4084 GATGAGAAATGGGGATTAATCTCTCAGTGGGGAAACGCACTTGTGCAATAGTAGAC 4143
Db 4165 GATGAGAAATGGGGATTAATCTCTCAGTGGGGAAACGCACTTGTGCAATAGTAGAC 4224
QY 4144 CCTGCTCGCCACTGTGAATCTGATTTAGATGAAGCACAAGCTGCCATGGACACAGA 4203
Db 4225 CCTGCTCGCCACTGTGAATCTGATTTAGATGAAGCACAAGCTGCCATGGACACAGA 4284
QY 4204 GACAGACTATGATTAAGAGACATCCGAGAGAGATTTGACAGACTGATACCATGCTGAC 4263
Db 4285 GACAGACTATGATTAAGAGACATCCGAGAGAGATTTGACAGACTGATACCATGCTGAC 4344
QY 4264 CATTTGCCATTCGCTGACACAGGTTCTAGGCTCCGATGAGATTAAGTGTGGCCACAGG 4323
Db 4345 CATTTGCCATTCGCTGACACAGGTTCTAGGCTCCGATGAGATTAAGTGTGGCCACAGG 4404
QY 4324 ACAAGTGGTGGAGTTGACACCCCATCGGCTCTCTGTCCAAACGACAGATTCGGATTTCA 4383
Db 4405 ACAAGTGGTGGAGTTGACACCCCATCGGCTCTCTGTCCAAACGACAGATTCGGATTTCA 4464
QY 4384 TGGCATGTTTGGCTGCGAGAGAAACAAGTGTGTCAGAGGGCTGACTCTCCCTTTGA 4443
Db 4465 TGGCATGTTTGGCTGCGAGAGAAACAAGTGTGTCAGAGGGCTGACTCTCCCTTTGA 4524
QY 4444 CGAAGTCTCTTTCTTTTAAAGCATTCGATTCCTGCTGGGGGGGGCCCTTCATCGCG 4503
Db 4525 CGAAGTCTCTTTCTTTTAAAGCATTCGATTCCTGCTGGGGGGGGCCCTTCATCGCG 4583
QY 4504 TCTCTCTACCGAAACCTTGCCTTTCTCGATTTATCTTCGACAGACATTCGGAATTTGG 4563
Db 4584 TCTCTCTACCGAAACCTTGCCTTTCTCGATTTATCTTCGACAGACATTCGGAATTTGG 4643
QY 4564 CTTGAGTCTTACCTTTTGGGAGAGTCAATTTGATTTGATTTATTCATATTTCA 4623
Db 4644 CTTGAGTCTTACCTTTTGGGAGAGTCAATTTGATTTGATTTATTCATATTTCA 4703
```


QY	2772	AAGAACCGGGAAACACCACCTGTCATCTCGAGGGAGACAGAACCTCGGTGAGTGCACAGATGA	2821
Db	2776	AAGAAACCGGGAACACCACCTGTCATCTCGAGGGAGACAGAACCTCGGTGAGTGCACAGATGA	2835
QY	2832	AGGACAACTCCTAATATGACAGTAATATGACAGATCATACGCCCTCTCCATGGGACATATGC	2891
Db	2836	AGGACAACTCCTAATATGACAGTAATATGACAGATCATACGCCCTCTCCATGGGACATATGC	2895
QY	2892	TGATCTCGAAAGCCATTCCAGAGCTTGTCTTTGTCAAGGGCACGCTGCGAGCTTCTCCC	2951
Db	2896	TGATCTCGAAAGCCATTCCAGAGCTTGTCTTTGTCAAGGGCACGCTGCGAGCTTCTCCC	2955
QY	2952	GGCTGCATGACGAGCTTTTCCGAAGATCCTTCGAAAGCCCTATGAAAGTTTMTTGTACACGA	3011
Db	2956	GGCTGCATGACGAGCTTTTCCGAAGATCCTTCGAAAGATCCTTCGAAAGCCCTATGAAAGTTTMTTGTACACGA	3015
QY	3012	CCCCACGAGGAGGATTTCTCACAGGTTTTTCCAAAGACATGAGATGATGACGTGCGGC	3071
Db	3016	CCCCACGAGGAGGATTTCTCACAGGTTTTTCCAAAGACATGAGATGATGACGTGCGCGC	3075
QY	3072	TGCGGTTCCAGGGCCGAGATTTATCATCAGAACGTTATCCCTGCTGTCCTTGTGTGGGAA	3121
Db	3076	TGCGGTTCCAGGGCCGAGATTTATCATCAGAACGTTATCCCTGCTGTCCTTGTGTGGGAA	3135
QY	3132	TGATCGGAGAGTCTTCCCGTGGTTCTTTGTGGCAGTGGGGGCCCTTGTCATCTCTTTT	3191
Db	3136	TGATCGGAGAGTCTTCCCGTGGTTCTTTGTGGCAGTGGGGGCCCTTGTCATCTCTTTT	3195
QY	3192	CAGTCTCGACATTTGTCTCCAGGGTCTCATTCGCGAGCTGAAAGGCTCGACATATCA	3251
Db	3196	CAGTCTCGACATTTGTCTCCAGGGTCTCATTCGCGAGCTGAAAGGCTCGACATATCA	3255
QY	3252	CGCAGTCACTTCTCTCTCCACATACGTCACAGATACAGGGCCCTTGCCACATCAACG	3311
Db	3256	CGCAGTCACTTCTCTCTCCACATACGTCACAGATACAGGGCCCTTGCCACATCAACG	3315
QY	3312	CCTTACATTAAGGGCAGAGATTTCTGCACAGATACAGAGCTGCTGGATGACAACCAAG	3371
Db	3316	CCTTACATTAAGGGCAGAGATTTCTGCACAGATACAGAGCTGCTGGATGACAACCAAG	3375
QY	3372	CTCCCTTTTTTTTTGTATACGTGTGCATCGCGGTGGCTGTGGGCTGGACCTCAACA	3431
Db	3376	CTCCCTTTTTTTTTGTATACGTGTGCATCGCGGTGGCTGTGGGCTGGACCTCAACA	3435
QY	3432	GCATCGCCCTCATCCACCAACGCGGCTGATGATGCTTCTTATGACGCGGCAGATTCCCC	3491
Db	3436	GCATCGCCCTCATCCACCAACGCGGCTGATGATGCTTCTTATGACGCGGCAGATTCCCC	3495
QY	3492	CAGCGCTATGGGGGCTCGGCCATCTCTTATGCTGTCCAGTTAACGGGGGCTTCCAGTTTA	3551
Db	3496	CAGCGCTATGGGGGCTCGGCCATCTCTTATGCTGTCCAGTTTAACGGGGGCTTCCAGTTTA	3555
QY	3552	CGGTCACTGGGACTCTGAGACAGAAAGCTCGATTCACTCGGTGGAGAGAGATCAATCACT	3611
Db	3556	CGGTCACTGGGACTCTGAGACAGAAAGCTCGATTCACTCGGTGGAGAGAGATCAATCACT	3615
QY	3612	ACATTAAAGACTGTCTCTTGTGGAAGCACTGCGCAAGATTAAAGAACAGGCTCCCTCCCTG	3671
Db	3616	ACATTAAAGACTGTCTCTTGTGGAAGCACTGCGCAAGATTAAAGAACAGGCTCCCTCCCTG	3675
QY	3672	ACTGAGCCCAAGAGGAGAGAGTACCTTTTGAGAACCAAGATGAGATACCGAGAAAC	3731
Db	3676	ACTGAGCCCAAGAGGAGAGAGTACCTTTTGAGAACCAAGATGAGATACCGAGAAAC	3735
QY	3732	TCCCTCTCTGCTTAAGAAAGTATCTTTCACGATCAAACTTAAGAGAGATTTGGCATTTG	3791
Db	3736	TCCCTCTCTGCTTAAGAAAGTATCTTTCACGATCAAACTTAAGAGAGATTTGGCATTTG	3795
QY	3792	TGGGGCCGAGAGATCAGGGAATCTCTCGCTGGGGATGGCCCTTCCGTCGTGTGGAGGT	3851
Db	3796	TGGGGCCGAGAGATCAGGGAATCTCTCTCGCTGGGGATGGCCCTTCCGTCGTGTGGAGGT	3855

QY	3852	TATCTGGAGCGCTGCATCAAGATTTGATGAGAGGAGATTCAGTATATGAGCTTCCGACAC	3911
Db	3856	TATCTGGAGCGCTGCATCAAGATTTGATGAGAGGAGATTCAGTATATGAGCTTCCGACAC	3915
QY	3912	TCCGAGCAAACTCTCTATCATTCCTCTCAAGACCGGTCCTTTGACAGTGACCTGCAGAT	3971
Db	3916	TCCGAGCAAACTCTCTATCATTCCTCTCAAGAGCGGTCCTTTGACAGTGACCTGCAGAT	3975
QY	3972	CAAAATTTGGACCCCTTCAACCAAGTACACTGTAAGACACAGATTTGGATGCCCTGGANAGGA	4031
Db	3976	CAAAATTTGGACCCCTTCAACCAAGTACACTGTAAGACACAGATTTGGATGCCCTGGANAGGA	4035
QY	4032	CACACATGAAGAAGATGTTTGTCTCAGCTACTGTAAGACACAGATTTGGATGCCCTGGANAGGA	4091
Db	4036	CACACATGAAGAAGATGTTTGTCTCAGCTACTGTAAGACACAGATTTGGATGCCCTGGANAGGA	4095
QY	4092	ATGGGGATTAACCTTCTCACTGGGGGGAACCGGACGCTCTTGTGATAGCTAGAGCCCTGCTCC	4151
Db	4096	ATGGGGATTAACCTTCTCACTGGGGGGAACCGGACGCTCTTGTGATAGCTAGAGCCCTGCTCC	4155
QY	4152	GCCACTGTAAATTCGTATTTTATGATGAGCCACACAGCTGCCATGGACACAGACAGACT	4211
Db	4156	GCCACTGTAAATTCGTATTTTATGATGAGCCACACAGCTGCCATGGACACAGACAGACT	4215
QY	4212	TATTCATTCAGAGACCATCCGAGAAAGATTTGGACATGTACATGCTGACCAATTTGCC	4271
Db	4216	TATTCATTCAGAGACCATCCGAGAAAGATTTGGACATGTACATGCTGACCAATTTGCC	4275
QY	4272	ATCGCCTGCACACGGTGTCTAGAGCTCCGATAGGATTTATGCTCTGGCCAGGAGACAGGTGG	4331
Db	4276	ATCGCCTGCACACGGTGTCTAGAGCTCCGATAGGATTTATGCTCTGGCCAGGAGACAGGTGG	4335
QY	4332	TGAGATTTTGACACCCCATTCGCTCTTCTGTCACACAGACATTTCCGATTTATGTCATGT	4391
Db	4336	TGAGATTTTGACACCCCATTCGCTCTTCTGTCACACAGACATTTCCGATTTATGTCATGT	4395
QY	4392	TTGCTGCTGCAGAGAACAAAGTGCCTGTCAGAGGCTGCACCTCCCTGTACACAAAGTCT	4451
Db	4396	TTGCTGCTGCAGAGAACAAAGTGCCTGTCAGAGGCTGCACCTCCCTGTACACAAAGTCT	4455
QY	4452	CTTTTCTTTAGAGCAATTCGCCATTCCTGCTGCGGGGCGGGCCCTTCATCGGCTCCTCTA	4511
Db	4456	CTTTTCTTTAGAGCAATTCGCCATTCCTGCTGCGGGGCGGGCCCTTCATCGGCTCCTCTA	4514
QY	4512	CCGAAACCTTGCCCTTCTCGATTTTATCTTTTCGACAGACAGTTCGGATTTGGCTTGTG	4571
Db	4515	CCGAAACCTTGCCCTTCTCGATTTTATCTTTTCGACAGACAGTTCGGATTTGGCTTGTG	4574
QY	4572	TTTTCATTTTAGGAGAGATCATTTTATATTTATTTATTTATTCATATTCATATGTAACA	4631
Db	4575	TTTTCATTTTAGGAGAGATCATTTTATATTTATTTATTTATTTATTCATATTCATATGTAACA	4634
QY	4632	AAATTTAGCTTTTGTCTCTTAATTCGACCTCTAAAGGTCAGGGAACGCTTATTTAATTTG	4691
Db	4635	AAATTTAGCTTTTGTCTCTTAATTCGACCTCTAAAGGTCAGGGAACGCTTATTTAATTTG	4694
QY	4692	TATCAGAGCGCTATTAATGAAGCTTTATACGTGTAGCTATATCTATATTAATTTCTGTACA	4751
Db	4695	TATCAGAGCGCTATTAATGAAGCTTTATACGTGTAGCTATATCTATATTAATTTCTGTACA	4754
QY	4752	TAGCTATATTTACAGTGAAGAAATGTAGCTGTTTATTTATATTTAAATTAAGCACTGTGC	4811
Db	4755	TAGCTATATTTACAGTGAAGAAATGTAGCTGTTTATTTATATTTAAATTAAGCACTGTGC	4814
QY	4812	TAAAAAAA 4819	
Db	4815	TAAATACA 4822	
RESULT	5		
LOCUS	AR091520	4781 bp	DNA
DEFINITION	Sequence 1 from patent US 5994130.		
		PAT	07-SEP-2000


```
|||||
Db 556 GCACAAAGCTGAGAGATGTGGCAAGAAAGCTGAATGAAGTTGGCCAGACGCTGTT 615
OY 612 CCTGCGAAGGTTGTGTGATCTTCTGCCACACAGGCTCATCTGTCCATCGTGTCC 671
Db 616 CCGTCGAAAGGTTGTGTGATCTTCTGCCACACAGGCTCATCTGTCCATCGTGTCC 675
OY 672 TGAATGATCAGCAGCTGGCTGGCTTCACTGAGACAGCCCTTCATGGTGAACACCTCTTGG 731
Db 676 TGAATGATCAGCAGCTGGCTGGCTTCACTGAGACAGCCCTTCATGGTGAACACCTCTTGG 735
OY 732 AGTATACCCAGGACAGAGTCTAACCCTGCACTACAGCTTGTGTATGTGCTGGCCCTCC 791
Db 736 AGTATACCCAGGACAGAGTCTAACCCTGCACTACAGCTTGTGTATGTGCTGGCCCTCC 795
OY 792 TCCGACGAGAAATGCTGGGCTTGTGGCTGTGCACTACCTGGGCAATGAAATACCGAA 851
Db 796 TCCGACGAGAAATGCTGGGCTTGTGGCTGTGCACTACCTGGGCAATGAAATACCGAA 855
OY 852 CCGGTGTCGGCTTGGGGGGCCATCCTAACCATGGCATTTAAGAAGATCCTTAAGTTAA 911
Db 856 CCGGTGTCGGCTTGGGGGGCCATCCTAACCATGGCATTTAAGAAGATCCTTAAGTTAA 915
OY 912 AGAATATTAAAGAGAAATCCCTGGGTGAGCTCATCAACATTTGCTCCACGATGGGCA 971
Db 916 AGAATATTAAAGAGAAATCCCTGGGTGAGCTCATCAACATTTGCTCCACGATGGGCA 975
OY 972 GAATGTTTGAAGACAGACCGCTTGGCAGCTGCTGGAGAGACCCGTTGTCATCT 1031
Db 976 GAATGTTTGAAGACAGACCGCTTGGCAGCTGCTGGAGAGACCCGTTGTCATCT 1035
OY 1032 TAGGATGATTTATTAATGTAATTAATCTGGAGACCAACAGGCTTCTGGATCAGCTGTT 1091
Db 1036 TAGGATGATTTATTAATGTAATTAATCTGGAGACCAACAGGCTTCTGGATCAGCTGTT 1095
OY 1092 TTATGCTCTTTTACCAGCAATGATGTTTGCATCAGGCTCAGACATATTTCAAGAGAA 1151
Db 1096 TTATGCTCTTTTACCAGCAATGATGTTTGCATCAGGCTCAGACATATTTCAAGAGAA 1155
OY 1152 AATGGTGGCCGACAGGATACGCTGCCAAGAAAGTAAGTAAGTTCTTTCATTTA 1211
Db 1156 AATGGTGGCCGACAGGATACGCTGCCAAGAAAGTAAGTAAGTTCTTTCATTTA 1215
OY 1212 AATTATCAAAATGATGATGCTGGGTCAAAAGCATTTTCTCAGAGTTTCAAGAAATCCGCG 1271
Db 1216 AATTATCAAAATGATGATGCTGGGTCAAAAGCATTTTCTCAGAGTTTCAAGAAATCCGCG 1275
OY 1272 AGAGAGAGCTGCGATATTGGAAGAAAGCCGGTACTTCCAGACATCACTGTGGTGG 1331
Db 1276 AGAGAGAGCTGCGATATTGGAAGAAAGCCGGTACTTCCAGAGCATCACTGTGGTGG 1335
OY 1332 CTCCTCATGTTGGTGTGATGTCACAGCGTGGTGAACCTTCTGTTTATATGACCCCTGGCT 1391
Db 1336 CTCCTCATGTTGGTGTGATGTCACAGCGTGGTGAACCTTCTGTTTATATGACCCCTGGCT 1395
OY 1392 TCGATCTGACAGCAGACAGGCTTTCACAGTGTGACAGTCTTCAATTCATGACTTTTG 1451
Db 1396 TCGATCTGACAGCAGACAGGCTTTCACAGTGTGACAGTCTTCAATTCATGACTTTTG 1455
OY 1452 CTTTGAAGATTAACCGTTTCACTAAGTCCCTCTCAGAAAGCTCAGTGGCTGTGACA 1511
Db 1456 CTTTGAAGATTAACCGTTTCACTAAGTCCCTCTCAGAAAGCTCAGTGGCTGTGACA 1515
OY 1512 GATTTAAGAGTTTGTGTTTATATGGAAGGTTTCACATGATTAAGAAACAAACAGCAGATC 1571
Db 1516 GATTTAAGAGTTTGTGTTTATATGGAAGGTTTCACATGATTAAGAAACAAACAGCAGATC 1575
OY 1572 CTCACATCAAGATAGAGATGAAAAATGCCACTTGGATGGAGATCTCCCACTCCAGATA 1631
Db 1576 CTCACATCAAGATAGAGATGAAAAATGCCACTTGGATGGAGATCTCCCACTCCAGATA 1635
OY 1632 TCCAGAACTCGCCCAAGCTGACCCCAAAATGAAAAAGACAAAGAGGCTTCCAGGGGCA 1691
|||||

Db 1636 TCCAGAACTCGCCCAAGCTGACCCCAAAATGAAAAAGACAAAGAGGCTTCCAGGGGCA 1695
OY 1692 AGAAAGAGAGGTGAGAGAGCTGCAGCCGCACTGACATCAAGCGGTGCTGGCAGACAGA 1751
Db 1696 AGAAAGAGAGGTGAGAGAGCTGCAGCCGCACTGACATCAAGCGGTGCTGGCAGACAGA 1755
OY 1752 AAGGCCACTCTCTCTGACAGTGAACAGCCGCCAAGTCCGAAAGAGAAAGCAAGC 1811
Db 1756 AAGGCCACTCTCTCTGACAGTGAACAGCCGCCAAGTCCGAAAGAGAAAGCAAGC 1815
OY 1812 ACATCCACTGGGCGCACTGGGCTTACAGAGAGACCTGCACAGCATGCTGAGATTC 1871
Db 1816 ACATCCACTGGGCGCACTGGGCTTACAGAGAGACCTGCACAGCATGCTGAGATTC 1875
OY 1872 AAGAGGTAAATGTTGGAATCTGCGGCACTGTGTGGAAGTGAAGAAACCTTCTCATATT 1931
Db 1876 AAGAGGTAAATGTTGGAATCTGCGGCACTGTGTGGAAGTGAAGAAACCTTCTCATATT 1935
OY 1932 CAGCCATTTTAAAGCCAGATGACCTTCTAGAGAGGCGACATTCGATGAGAACCTTCG 1991
Db 1936 CAGCCATTTTAAAGCCAGATGACCTTCTAGAGAGGCGACATTCGATGAGAACCTTCG 1995
OY 1992 CTTATGTGGCCAGAGCGCTGGATCTCAATGCTACCTGAGAGACAAACATCCTGTTG 2051
Db 1996 CTTATGTGGCCAGAGCGCTGGATCTCAATGCTACCTGAGAGACAAACATCCTGTTG 2055
OY 2052 GGAAGGAATATGATGAAGAAAGATCAACTCTGTGCTGAACAGCTGCTGAGAGGCTG 2111
Db 2056 GGAAGGAATATGATGAAGAAAGATCAACTCTGTGCTGAACAGCTGCTGAGAGGCTG 2115
OY 2112 ACCTGGCAATTTCTCCAGACAGCAGCTGACGAGATTTGGAGAGCGAGAGCCAACTGA 2171
Db 2116 ACCTGGCAATTTCTCCAGACAGCAGCTGACGAGATTTGGAGAGCGAGAGCCAACTGA 2175
OY 2172 GCGGTGGGAGGCGCAGAGGATCAGCTTGGCCGGGCTGTGATGATGACAGAGACATCT 2231
Db 2176 GCGGTGGGAGGCGCAGAGGATCAGCTTGGCCGGGCTGTGATGATGACAGAGACATCT 2235
OY 2232 ACATCCTGAGACACCCCTCAGTGCCTTAGATGCCATGTGGGCAACCAATCTTCAATA 2291
Db 2236 ACATCCTGAGACACCCCTCAGTGCCTTAGATGCCATGTGGGCAACCAATCTTCAATA 2295
OY 2292 GTGCTATCCGGAACATCTCAAGTCCAAAGACAGTCTGTTGTTACCAACAGTTACGT 2351
Db 2296 GTGCTATCCGGAACATCTCAAGTCCAAAGACAGTCTGTTGTTACCAACAGTTACGT 2355
OY 2352 ACCTGTTGACCTGTATGAAGATCTTCATGAAGAGAGGCTGTATTACGGAAGAGGCA 2411
Db 2356 ACCTGTTGACCTGTATGAAGATCTTCATGAAGAGAGGCTGTATTACGGAAGAGGCA 2415
OY 2412 CCCATGAGAACTGATGAATTTAATGTTGACTATGCTACCAATTTTAAATACTGTTGC 2471
Db 2416 CCCATGAGAACTGATGAATTTAATGTTGACTATGCTACCAATTTTAAATACTGTTGC 2475
OY 2472 TGGGAGAGACACCGCAGTGTGATCAATTTCAAAAAGAGAAACAGTGGTTCCAGAGAA 2531
Db 2476 TGGGAGAGACACCGCAGTGTGATCAATTTCAAAAAGAGAAACAGTGGTTCCAGAGAA 2535
OY 2532 AGTCACAAAGACAGGCTCTTAAACAGATCAATTAAGAAAGGAAAAACAGTAAGCCAG 2591
Db 2536 AGTCACAAAGACAGGCTCTTAAACAGATCAATTAAGAAAGGAAAAACAGTAAGCCAG 2595
OY 2592 AGGAAGGAGAGCTGTGACGCTGGAAGAGAAAGGCGAGGTTCACTGGCCCTGCTAGTAT 2651
Db 2596 AGGAAGGAGAGCTGTGACGCTGGAAGAGAAAGGCGAGGTTCACTGGCCCTGCTAGTAT 2655
OY 2652 ATGCTGTACATCCAGGCTGCTGGGGGCCCTTGGCATTTCTGTTATATATGAGCCCTTT 2711
Db 2656 ATGCTGTACATCCAGGCTGCTGGGGGCCCTTGGCATTTCTGTTATATATGAGCCCTTT 2715
OY 2712 TCATGCTGAATGTAGGACAGACCGCTTTCAGACCTGCTGTGATTAAGTGAATCAAGC 2771
Db 2716 TCATGCTGAATGTAGGACAGACCGCTTTCAGACCTGCTGTGATTAAGTGAATCAAGC 2775
```

```
|||||
Db 1861 CATTTCGACATTTTAGCCAGATGACGCTTCTAGAGGGAGCATTCATTCAGTGAAC 1920
|||
Qy 1987 CTTTCGCTTATGTCGCCAGCAGAGCGCTGGATCTCTCAATGATCTACTCTGAGAGACAATCT 2046
|||
Db 1921 CTTTCGCTTATGTCGCCAGCAGAGCGCTGGATCTCTCAATGATCTACTCTGAGAGACAATCT 1980
|||
Qy 2047 GTTTCGGAAGGAATATGATGAAGAAATACAACTCTGTCTGAACAGCTGCTGCCAG 2106
|||
Db 1981 GTTTCGGAAGGAATATGATGAAGAAATACAACTCTGTCTGAACAGCTGCTGCCAG 2040
|||
Qy 2107 GCGTCGACCTGGCATCTCTTCCAGCAGCAGCCTGACGAGATTTGAGAGAGAGGCCAA 2166
|||
Db 2041 GCGTCGACCTGGCATCTCTTCCAGCAGCAGCCTGACGAGATTTGAGAGAGAGGCCAA 2100
|||
Qy 2167 CCGTACGCTGGGCGACGCCAGAGGATCAGCCCTTGCCGGGCTTGTATAGTCAGAG 2226
|||
Db 2101 CCGTACGCTGGGCGACGCCAGAGGATCAGCCCTTGCCGGGCTTGTATAGTCAGAG 2160
|||
Qy 2227 CATCTCATCTGGAGAGACCCCTCAGTGCCTAGATGCCATGTGGCAACCATCTT 2286
|||
Db 2161 CATCTCATCTGGAGAGACCCCTCAGTGCCTAGATGCCATGTGGCAACCATCTT 2220
|||
Qy 2287 CAATAGTCTATCCGGAACAATCTCAAGTCCAAAGACAGTTCTGTTGTACCCACAGTT 2346
|||
Db 2221 CAATAGTCTATCCGGAACAATCTCAAGTCCAAAGACAGTTCTGTTGTACCCACAGTT 2280
|||
Qy 2347 ACAGTACCTGCTGATCTGATGATGATGATCTTCATGAAGAGGCGTGTATACGGAAG 2406
|||
Db 2281 ACAGTACCTGCTGATCTGATGATGATGATCTTCATGAAGAGGCGTGTATACGGAAG 2340
|||
Qy 2407 AGGACCCATGAGGAATCTGATGAATTTAAATGGTACTATGCTACATTTTATTAACCT 2466
|||
Db 2341 AGGACCCATGAGGAATCTGATGAATTTAAATGGTACTATGCTACATTTTATTAACCT 2400
|||
Qy 2467 GTTTCGTCGAGAGACACCGCCAGTGGATCAATTCAAAAAGAAACCAAGTGGTTCACA 2526
|||
Db 2401 GTTTCGTCGAGAGACACCGCCAGTGGATCAATTCAAAAAGAAACCAAGTGGTTCACA 2460
|||
Qy 2527 GAAGAATCTCAAGACAAAGGCTCTTAAGAGATCAATTAAGAAAGAAAGAAAGAGTAA 2586
|||
Db 2461 GAAGAATCTCAAGACAAAGGCTCTTAAGAGATCAATTAAGAAAGAAAGAAAGAGTAA 2520
|||
Qy 2587 GCCAGAGGAAGGCGACTTGTGACGTGGAAGAGAAAGGCGAGGCTCAGGCGCTGCTC 2646
|||
Db 2521 GCCAGAGGAAGGCGACTTGTGACGTGGAAGAGAAAGGCGAGGCTCAGTGCCTGCTC 2580
|||
Qy 2647 AGTATATGCTGTCTACATCCAGGCTGCTGGGGGCCCTTGGCATTCCTGTTATTTATG 2706
|||
Db 2581 AGTATATGCTGTCTACATCCAGGCTGCTGGGGGCCCTTGGCATTCCTGTTATTTATG 2640
|||
Qy 2707 CATTTCATGCTGAATGATGAGGACACCGGCTTTCAGACCTGGTGTGATGATCTGAT 2766
|||
Db 2641 CATTTCATGCTGAATGATGAGGACACCGGCTTTCAGACCTGGTGTGATGATCTGAT 2700
|||
Qy 2767 CAAGCAAGGAAGCGGGAACCACTGTGATCGAGGGAAGAGACCTCGGAGTGAACAG 2826
|||
Db 2701 CAAGCAAGGAAGCGGGAACCACTGTGATCGAGGGAAGAGACCTCGGAGTGAACAG 2760
|||
Qy 2827 CATTGAAGACAATCTCATATGACAGTATATGCCAGCATCTACGCCCTCTCCATGAGCAGT 2886
|||
Db 2761 CATTGAAGACAATCTCATATGACAGTATATGCCAGCATCTACGCCCTCTCCATGAGCAGT 2820
|||
Qy 2887 CATTGCTATCTGAAAGCAATGAGAGATGCTTGTGTAAGGGGACGCTGCGACAGCTTC 2946
|||
Db 2821 CATTGCTATCTGAAAGCAATGAGAGATGCTTGTGTAAGGGGACGCTGCGACAGCTTC 2880
|||
Qy 2947 CTCGCGGCTCATACAGAGCTTTTCCGAAGGATCTTCGAAGCCCTATGAAGTTTGA 3006
|||
Db 2881 CTCGCGGCTCATACAGAGCTTTTCCGAAGGATCTTCGAAGCCCTATGAAGTTTGA 2940
|||
Qy 3007 CACGACCCCAACAGGAGATCTCAACAGGTTTCCAAAGACATGAGTGAAGTGAAGCT 3066
|||
```

ACCESSION AR091520
VERSION AR091520.1 GI:10018275
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 4781)
TITLE Shyjan, A.
JOURNAL Multidrug resistance-associated polypeptide
Patent: US 5994130-A 1 30-NOV-1999;
FEATURES Location/Qualifiers
source 1..4781
BASE COUNT 1217 a 1161 c 1232 g 1165 t 6 others
ORIGIN

Query Match 98.58; Score 4775.4; DB 9; Length 4781;
Best Local Similarity 99.88; Pred. No. 0;
Matches 4773; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

67 TGATGTGAACCTAAGCTGTGAGCCCTGGAACCTCCACATCAGAGAGATGAAGATAT 126
1 TGATGTGAACCTAAGCTGTGAGCCCTGGAACCTCCACATCAGAGAGATGAAGATAT 60
127 CGACATAGGAAGAGATATCATCCAGCTCTGGTATAGAGTGTGAGGAGAGAAC 186
61 CGACATAGGAAGAGATATCATCCAGCTCTGGTATAGAGTGTGAGGAGAGAAC 120
187 CAGACTTCTGGGAGCAGACAGACCGTGAAGATTCCAGTTCCAGAGACTCGACGTT 246
121 CAGACTTCTGGGAGCAGACAGACCGTGAAGATTCCAGTTCCAGAGACTCGACGTT 180
247 GGAATGCCAAGATGCTTGAAGACAGACCGCGAGCGGAGGCGCTCTCTTGAATGCTC 306
181 GGAATGCCAAGATGCTTGAAGACAGACCGCGAGCGGAGGCGCTCTCTTGAATGCTC 240
307 CATGCATCTCAGCTCAGAAATCTGGATGAGAGACATCCCAAGGGAAGTACCATATG 366
241 CATGCATCTCAGCTCAGAAATCTGGATGAGAGACATCCCAAGGGAAGTACCATATG 300
367 CTGAGTCTCTGAAGCCCATCCGAGTACTTGCACACACAGACCCCATGAGCATGTC 426
301 CTGAGTCTCTGAAGCCCATCCGAGTACTTGCACACACAGACCCCATGAGCATGTC 360
427 TGGGTTTTTCTGATGACTTTTCTGGGCTTCTCTGCGCCGCTGGGCCACAA 486
361 TGGGTTTTTCTGATGACTTTTCTGGGCTTCTCTGCGCCGCTGGGCCACAA 420
487 GAAGGGGAGCTCTCAATGAGAAGAGTGTGCTCTGTCCAAAGCAGAGTCTTGACGT 546
421 GAAGGGGAGCTCTCAATGAGAAGAGTGTGCTCTGTCCAAAGCAGAGTCTTGACGT 480
547 GAATCTGAGAAGACTAGAGACTGTGGCAAGAGAGCTGAATGAAGTTGGGCCAGACC 606
481 GAATCTGAGAAGACTAGAGACTGTGGCAAGAGAGCTGAATGAAGTTGGGCCAGACC 540
607 TGGTCCCTGGGAAGGGTGTGTGATCTTCTGCGGCACAGGCTCACTCTGCATGCT 666
541 TGGTCCCTGGGAAGGGTGTGTGATCTTCTGCGGCACAGGCTCACTCTGCATGCT 600
667 GTGCGTATGATCAAGCAGCTGCTGCTTCAAGTGAAGCAGCTTCAATGGAACACCT 726
601 GTGCGTATGATCAAGCAGCTGCTGCTTCAAGTGAAGCAGCTTCAATGGAACACCT 660
727 CTGAGCATATACCCAGGACAGAGTCTAACCCTGCAGTACAGCTTGTGTTAGTCTGG 786
661 CTGAGCATATACCCAGGACAGAGTCTAACCCTGCAGTACAGCTTGTGTTAGTCTGG 720
787 CCTCTCTCTGAGGAATCGTGGCGTCTTGGTCTGCTGACTGACTTGGGATGAAATTA 846
721 CCTCTCTCTGAGGAATCGTGGCGTCTTGGTCTGCTGACTGACTTGGGATGAAATTA 780

847 CCGAACCGGTGCTCCGCTTGGCGGGGGCCATCTTAACATGGCATTTAAGAGATCTTAA 906
781 CCGAACCGGTGCTCCGCTTGGCGGGGGCCATCTTAACATGGCATTTAAGAGATCTTAA 840
907 GTTAAAGAACATTTAAAGGAAATCCCTGGGTAGCTCATCAACATTTCTCCAGCATGG 966
841 GTTAAAGAACATTTAAAGGAAATCCCTGGGTAGCTCATCAACATTTCTCCAGCATGG 900
967 GCAGAGATGTTTGGAGAGCAGCCGTTGGCAGCTGCTGGCTGAGAGACCGTGTGTC 1026
901 GCAGAGATGTTTGGAGAGCAGCCGTTGGCAGCTGCTGGCTGAGAGACCGTGTGTC 960
1027 CATCTTAGCATGATTTATTAATGTAATTTCTGGGACCAACAGGCTTCTGGGATCAGC 1086
961 CATCTTAGCATGATTTATTAATGTAATTTCTGGGACCAACAGGCTTCTGGGATCAGC 1020
1087 TGTTTTATCCTCTTTTACCCAGCATATGTTTTCATCAGCGCTCAGCATATTTAG 1146
1021 TGTTTTATCCTCTTTTACCCAGCATATGTTTTCATCAGCGCTCAGCATATTTAG 1080
1147 GAGAAATGCGTGGCCGCGCAGGATGAACGTGTCCAGAAAGTGAAGTCTTACTTTA 1206
1081 GAGAAATGCGTGGCCGCGCAGGATGAACGTGTCCAGAAAGTGAAGTCTTACTTTA 1140
1207 CATTAATTTTACAAATATGATGCTTGGGTCAAAAGCATTTTCTCAGAGTGTTCAGAAAT 1266
1141 CATTAATTTTACAAATATGATGCTTGGGTCAAAAGCATTTTCTCAGAGTGTTCAGAAAT 1200
1267 CCGGAGAGAGAGCTCGGATTTTGGAAAAAGCCGGGTACTTCCAGAGCATCAGTGTGG 1326
1201 CCGGAGAGAGAGCTCGGATTTTGGAAAAAGCCGGGTACTTCCAGAGCATCAGTGTGG 1260
1327 TGTGCTCCCATTTGGTGGTGAATGATGCCAGCGTGTGACCTCTCTGTTCAATGACCT 1386
1261 TGTGCTCCCATTTGGTGGTGAATGATGCCAGCGTGTGACCTCTCTGTTCAATGACCT 1320
1387 GGGCTTCATCTGACAGACGACAGAGCTTTCACAGTGTGACAGTCTTCCATGAC 1446
1321 GGGCTTCATCTGACAGACGACAGAGCTTTCACAGTGTGACAGTCTTCCATGAC 1380
1447 TTTTCTTTGAAGTAACGCTTTTCAGTAAGCCCTTCAGAGAGCTCAGTGGCTGT 1506
1381 TTTTCTTTGAAGTAACGCTTTTCAGTAAGCCCTTCAGAGAGCTCAGTGGCTGT 1440
1507 TGACAGATTTAAGATTTGTTTCTAATGGAAGAGTTCACATGATAAAGCAACACAGC 1566
1441 TGACAGATTTAAGATTTGTTTCTAATGGAAGAGTTCACATGATAAAGCAACACAGC 1500
1567 CAGTCTCATCATCAAGATGAGATGAAAAATGCCACTTGGCATGGACTCTCCCACTC 1626
1501 CAGTCTCATCATCAAGATGAGATGAAAAATGCCACTTGGCATGGACTCTCCCACTC 1560
1627 CAGTATCCAGAACTGCCCCAGCTACCCCCCAAAATGAAAAAGCAAGAGGCTTCCAG 1686
1561 CAGTATCCAGAACTGCCCCAGCTACCCCCCAAAATGAAAAAGCAAGAGGCTTCCAG 1620
1687 GGGCAAGAAAGAGAGTGAAGCAGCTGAGCGACCTGACGCTCAGGCGGTGCTGGCAGA 1746
1621 GGGCAAGAAAGAGAGTGAAGCAGCTGAGCGACCTGACGCTCAGGCGGTGCTGGCAGA 1680
1747 GCAGAAAGCCACCTCTCTCTGAGACGTAGCAGCGGCCACAGTCCCGAAGGAAGAAAG 1806
1681 GCAGAAAGCCACCTCTCTCTGAGACGTAGCAGCGGCCACAGTCCCGAAGGAAGAAAG 1740
1807 CAAACATATCCACTGGGACACCTGCGCTTACAGAGACACTGCACAGCATGATCTGGA 1866
1741 CAAACATATCCACTGGGACACCTGCGCTTACAGAGACACTGCACAGCATGATCTGGA 1800
1867 GATCCAAAGAGGTAACGTGTTGAAATCTGGGCGCAGTGTGGGAATGGAACCTCTCT 1926
1801 GATCCAAAGAGGTAACGTGTTGAAATCTGGGCGCAGTGTGGGAATGGAACCTCTCT 1860
1927 CATTTACGCCATTTTAGCGCATGACGCTTCTAGAGGGCAGCATTTGCAATCAGTGAAC 1986

Pharmacology Division; Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045,
Japan (E-mail:tsuzuki@gen2.res.ncc.go.jp,
Tel:81-3-3542-2511(ex.4451), Fax:81-3-3542-1886)
Location/Qualifiers

1..5728
/organism="Homo sapiens"
/db_xref="taxon:9606"
/gene="MRP5"
53..4366
/gene="MRP5"
53..4366
/codon_start=1
/protein_id="BAA76608.1"
/db_xref="GI:4587083"

/translation="MKDIDICKEYIIPSPGTRSVNERISTGTHHREDSEKRRRL
ECQDALLETAAAREGLISLDSMSQRLIDEHPKGRVHGLSKLPKRTSKHOPVD
NAGLESCMTFMSLSLARVARKKGLSMEDWVSLKHSVNCRLRLMOELNVEY
GPDASLRVYVIECPRLILSYCLMTOLAGFSGPAPVYKHLLEYQAESNLQYS
LLIVGLITFEIVRSMSLITMALNYPVGRVRLGLILMAPEKILIKTEKSGEL
INICNDQGRPEAAVSGSLAGPVAVILGHYVIVITLGPGLGSAVITLFPYAM
FASRLTAYFRKCVAAVDEROAKNEVLTIKETIMAVAFVSQVRIEERIL
EKAGYFOSITVGVAPIVVIVIASVYFVSHMTLGPDLTAQAFTVTVENSMTFALKT
PEYSLSLEASVADRFKSLFMEVHMKKPPHRIKEMKNTLMDSSHSIQN
SPKLPKKMKDKRASRGKEKERVROLDREHOAVLAEKQHLILDSDEPSEEBGKH
THGLRLORTLSIDLEIOBKIVIGIGSVSGSKTSLISAILGOMTLEGS IATSGT
FAYVAQMIILNATLRDNIILKEKEDERYNSVYNSCLRPDLATLPSDLTEIGERG
ANLGGORISLARALYSDSITLIDPLSLDLDHVNHLFNSAIRHLKSTVLFV
THOLOIVADCEDEIVFEMKEGCIYTERGTHELNLDNDAVITFNMLLIGTPPEVINSK
ETSGSKKSPDKGPKTGSVKKERKAVKPEEQOLVLEEGOGSVPSVGVYIOAGGP
LAFVIMLPMNVGSPAFSVMWLSYMKOGSGNTVYRGNETSVDSKMDPHOYV
ASTVLSMAVMILIKAIRGVYVKGTLTAASSSKHDELRRILSRSPKFPDTPPRL
NRFSDMDVDRLEPFOAEMFIONIIIVPEYGMHAGYFPMVLAVGLVILFVSLHI
VSRVILRELKRLDNTQSPFLSHITSIQGLATIHAYKGGEFFRLRDELDNDAPF
FLETCAMWLVARDLISIALITTTGLMIVLMHGOIPAYVGLAISVAVQLGLEFT
VRLASETEAREFVSERINHYIKTSLLEAPARKNAPSDVPOBEVFEENAEMRYE
NLPLVKVSEFTIKERKIGIVRGTSKSSIGMALFRLVELSGCIKIDGVRISDIG
LADRSKLSIIPQEVLPFSGVYRSLNLDPEFNTOEDIVALESTHMKECIAOLPLKE
SEVMNGNFSVGFOLLICIRALLRHCKITLIDDEATAMDEFTDILIOETRFKAD
CTMLNIAHRLHRTVIGSDRIMVLAQGVYEPDTPSVLLSNDSSRFYAMFAAENKAVK
G"

FEATURES
source
gene
CDS

4147 GCTCCGCCACTGTAGATTCTGATTTTATGATGAGCCACAGCTGCCATGACACAGAGAC 4206
|||||
4081 GCTCCGCCACTGTAGATTCTGATTTTATGATGAGCCACAGCTGCCATGACACAGAGAC 4140
4207 AACACTTATGATTCAAGAGACCATCCGAGACATTTGAGAGCTGACATGCTGACAT 4266
|||||
4141 AGACTTATGATTCAAGAGACCATCCGAGACATTTGAGAGCTGACATGCTGACAT 4200
4267 TGCCTATGCTGACACAGCTTCTAGGCTCGATAGATTATGCTGCTGGCCAGGAGACA 4326
|||||
4201 TCCCATGCGCTGACACAGCTTCTAGGCTCGATAGATTATGCTGCTGGCCAGGAGACA 4260
4327 GGTGGTGGAGTTTGACACCCCATCGCTCTCTGCTCCACAGACAGTCCCGATCTATCC 4386
|||||
4261 GGTTGGGAGTTTGACACCCCATCGCTCTCTGCTCCACAGACAGTCCCGATCTATCC 4320
4387 CATGTTTGTGCTGACAGACAGAGTCCGCTGCAAGGCTGACCTCCCTGCTTGACAGA 4446
|||||
4321 CATGTTTGTGCTGACAGACAGAGTCCGCTGCAAGGCTGACCTCCCTGCTTGACAGA 4380
4447 AGCTCTTTTCTTTAGAGAGATTGCCATTCCTGCTGGGGGGGGCCCTTCATCCGCTCC 4506
|||||
4381 AGCTCTTTTCTTTAGAGAGATTGCCATTCCTGCTGGGGGGGGCCCTTCATCCGCTCC 4440
4507 TCCCTACCGAAACCTTGCCTTCTCCATTTATCTTTCCGACAGACATTCGGGATTGGCTT 4566
|||||
4441 TCCCTACCGAAACCTTGCCTTCTCCATTTATCTTTCCGACAGACATTCGGGATTGGCTT 4500
4567 GGTGTTTCTACTTTATGAGAGAGATCATTTTATGATTATGATTATTCATATTCATGCT 4626
|||||
4501 GGTGTTTCTACTTTATGAGAGAGATCATTTTATGATTATGATTATTCATATTCATGCT 4560
4627 AAACAAATTTAGTTTGTCTTTAATTCACACTTAAAGGTTCAAGGAACCGTTATTAAT 4686
|||||
4561 AAACAAATTTAGTTTGTCTTTAATTCACACTTAAAGGTTCAAGGAACCGTTATTAAT 4620
4687 AATTGATAGAGGCTTAAATGAGGCTTATAGCGTATAGCTATATCTATATATATTTCT 4746
|||||
4621 AATTGATAGAGGCTTAAATGAGGCTTATAGCGTATAGCTATATCTATATATATTTCT 4680
4747 GTACATAGCTTATTTACAGTGAATGTAAGCTTTATTTATTTATTTAAATAAGACAC 4806
4681 GTACATAGCTTATTTACAGTGAATGTAAGCTTTATTTATTTATTTAAATAAGACAC 4740
4807 TGTGCTAAAAAAGGCGCGCCG 4847
4741 TGTGCTAAAAAAGGCGCGCCG 4781

RESULT 6
AB019002 5728 bp mRNA PRI 01-MAR-2000
DEFINITION Homo sapiens MRP5 mRNA, complete cds.
ACCESSION AB019002
VERSION AB019002.1 GI:4587082
KEYWORDS
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Suzuki,T., Sasaki,H., Kuh,H.J., Agui,M., Tatsumi,Y., Tanabe,S.,
Terada,M., Saijo,N. and Nishio,K.
TITLE Detailed structural analysis on both human MRP5 and mouse mrp5
transcripts
JOURNAL Gene 242 (1-2), 167-173 (2000)
MEDLINE 2 (bases 1 to 5728)
REFERENCE Suzuki,T., Kuh,H. and Nishio,K.
AUTHORS Direct Submission
TITLE Submitted (20-OCT-1998) to the DBJ/EMBL/Genbank databases.
JOURNAL Toshitiro Suzuki, National Cancer Center Research Institute,

Query Match 97.5%; Score 4728; DB 85; Length 5728;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 4745; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

polyA.site
BASE COUNT 1379 a 1415 c 1472 g 1462 t
ORIGIN

64 TTCTGATGTGAACCTAACAGTCTGTGAGCCCTGGAACTCTCCACTCGAGAGAAATGAAGA 123
|||||
1 TTCTGATGTGAACCTAACAGTCTGTGAGCCCTGGAACTCTCCACTCGAGAGAAATGAAGA 60
124 TATCGACATAGGAAAAGATATATATCATCCCGAGTCTGGGTATAGAAAGTGTGAGAGAG 183
61 TATCGACATAGGAAAAGATATATATCATCCCGAGTCTGGGTATAGAAAGTGTGAGAGAG 120
184 AACCAACACTTGTGGAGCCACAGAGACCGTGAAGATTTCCAAAGTTTCAGAGAACTTGACC 243
121 AACCAACACTTGTGGAGCCACAGAGACCGTGAAGATTTCCAAAGTTTCAGAGAACTTGACC 180
244 GTTGAATGCCAAGATGCTTTGGAACACAGACCGCAGGCGCTCTCTCTTATATGC 303
181 GTTGAATGCCAAGATGCTTTGGAACACAGACCGCAGGCGCTCTCTCTTATATGC 240
304 CTCGATGATTCACATCGACATGAGATCTCGATGAGAGATCCCAAGGAAAGTACCATCA 363
241 CTCGATGATTCACATCGACATGAGATCTCGATGAGAGATCCCAAGGAAAGTACCATCA 300
364 TGGCTTGAAGTGTCTGTAAGCCCATCCGAGCTACTCTTCCAAACACACACACCCAGTGGACA 423
301 TGGCTTGAAGTGTCTGTAAGCCCATCCGAGCTACTCTTCCAAACACACACACCCAGTGGACA 360

|||||
Db 2521 AAGCCAGAGGAGGCGAGCTGTGCACGTGGAAAGAAAGGCGAGGTTGAGTCCCTG 2580
OY 2644 GTCACATATGATGTCATCATATCCAGAGCTGCTGGGGCCCCCTTGAGATCTCGGTATAT 2703
Db 2591 GTCAATATATGATGTCATCATATCCAGAGCTGCTGGGGCCCCCTTGAGATCTCGGTATAT 2640
OY 2704 GGCCTTTTCATGCTGATGATGATGAGCAGCAGCCCTTCAGACATCTGCTGATGATGATG 2763
Db 2641 GGCCCTTTTCATGCTGATGATGATGAGCAGCAGCCCTTCAGACATCTGCTGATGATGATG 2700
OY 2764 GATCAAGCAAGAGAGGAGGAGAACACACACCTGTGACTCGAGGAGAACGAGACCTCGTGA 2823
Db 2701 GATCAAGCAAGAGAGGAGGAGAACACACACCTGTGACTCGAGGAGAACGAGACCTCGTGA 2760
OY 2824 CAGCATGAAGCAATCTCTCATATGACATGATGACACATCTACGCCCTCTCATGAGC 2883
Db 2761 CAGCATGAAGCAATCTCTCATATGACATGATGACACATCTACGCCCTCTCATGAGC 2820
OY 2884 AGTCATGCTGATCCCTGAAGGCATTCGAGGAGTTGCTTTGCAAGGGGACGCTGCGAGC 2943
Db 2821 AGTCATGCTGATCCCTGAAGGCATTCGAGGAGTTGCTTTGCAAGGGGACGCTGCGAGC 2880
OY 2944 TTCCCTCCCGGCTGATGAGCAGAGCTTTCCGAGAGATCCCTTCGAGACCCCTATGAGTTT 3003
Db 2881 TTCCCTCCCGGCTGATGAGCAGAGCTTTCCGAGAGATCCCTTCGAGACCCCTATGAGTTT 2940
OY 3004 TGACACGACCCCGACAGGAGAGATCTCAACAGTTTTCACAAAGACATGATGAGAGTTGA 3063
Db 2941 TGACACGACCCCGACAGGAGAGATCTCAACAGTTTTCACAAAGACATGATGAGAGTTGA 3000
OY 3064 CGTGGGCTGGCGCTCCAGGCGGAGATGTTCAACAGAGTTATCCGAGTCTGCTG 3123
Db 3001 CGTGGGCTGGCGCTCCAGGCGGAGATGTTCAACAGAGTTATCCGAGTCTGCTG 3060
OY 3124 TGTGGGAATGATGAGCAGAGAGTCTCCCGTGTGTTCTTGTGAGCAGTGGGCCCTTGTGAT 3183
Db 3061 TGTGGGAATGATGAGCAGAGAGTCTCCCGTGTGTTCTTGTGAGCAGTGGGCCCTTGTGAT 3120
OY 3184 CCTCTTTTACGTCTGACACATGTTCTCCAGAGGCTCTGATGCGGAGCTGAGTGA 3243
Db 3121 CCTCTTTTACGTCTGACACATGTTCTCCAGAGGCTCTGATGCGGAGCTGAGTGA 3180
OY 3244 CAATATCAAGCAGTACACCTTCTCTCCACATCACTGACATACAGGGCTGGCCAC 3303
Db 3181 CAATATCAAGCAGTACACCTTCTCTCCACATCACTGACATACAGGGCTGGCCAC 3240
OY 3304 CATCAGCAGCTACATTAAGGAGGAGAGTTTCTGCACAGATACAGAGACTCTGATGA 3363
Db 3241 CATCAGCAGCTACATTAAGGAGGAGAGTTTCTGCACAGATACAGAGACTCTGATGA 3300
OY 3364 CAACCAAGCTCTTTTTTTTTTTTTTGTAACTGTGCGATGCGGCTGAGTGGCTGGA 3423
Db 3301 CAACCAAGCTCTTTTTTTTTTTTTTGTAACTGTGCGATGCGGCTGAGTGGCTGGA 3360
OY 3424 CCTCATACGATCGGCTCATATCCACACAGGGGCGATGATGCTTTATGACAGGGGA 3483
Db 3361 CCTCATACGATCGGCTCATATCCACACAGGGGCGATGATGCTTTATGACAGGGGA 3420
OY 3484 GATTCGCCAGCCTATGCGGCTGCGCATCTTATGCTGTCCAGTTAACGGGCTGT 3543
Db 3421 GATTCGCCAGCCTATGCGGCTGCGCATCTTATGCTGTCCAGTTAACGGGCTGT 3480
OY 3544 CCAGTTTACGCTGACAGCTGCGCATGTGAGACAGAGCTGATTCACCTCGGTGAGAGAT 3603
Db 3481 CCAGTTTACGCTGACAGCTGCGCATGTGAGACAGAGCTGATTCACCTCGGTGAGAGAT 3540
OY 3604 CAATATACATTAAGACATGCTGCTGGAAGCAGCTGCAATTAAGAAAGAGCTCC 3663
Db 3541 CAATATACATTAAGACATGCTGCTGGAAGCAGCTGCAATTAAGAAAGAGCTCC 3600
OY 3664 CTCCCTGACTGGCCCCAGAGGAGAGTGAACCTTGAAGACGAGATGAGTACCG 3723
|||||

Db 3601 CTCCCTGACTGGCCCCAGAGGAGAGTGAACCTTTGAGAACGACGAGATGAGGTACCG 3660
OY 3724 AGAAACCTCCCTCGTCTGCTTAAGAAAGTATCCCTACGATCAAACTTAAGAGAGAT 3783
Db 3661 AGAAACCTCCCTCGTCTGCTTAAGAAAGTATCCCTACGATCAAACTTAAGAGAGAT 3720
OY 3784 TGGCATTTGGGGGCGAGAGATGACAGGAAAGCTCGCTGGGGATGGCCCTTCCGCT 3843
Db 3721 TGGCATTTGGGGGCGAGAGATGACAGGAAAGCTCGCTGGGGATGGCCCTTCCGCT 3780
OY 3844 GGTGAGTTATCTGAGGCTGATCAAGATGATGAGAGTGAATGATGATTTGGCCT 3903
Db 3781 GGTGAGTTATCTGAGGCTGATCAAGATGATGAGAGTGAATGATGATTTGGCCT 3840
OY 3904 TGCCCACTCCCAACCAAACTCTATATCTTCAAGACCGGGGCTGTGAGTGGAC 3963
Db 3841 TGCCCACTCCCAACCAAACTCTATATCTTCAAGACCGGGGCTGTGAGTGGAC 3900
OY 3964 TGTGATCAATTTTGAACCCCTTCACACAGTACAGTGAAGACAGATTTGGATGCCCT 4023
Db 3901 TGTGATCAATTTTGAACCCCTTCACACAGTACAGTGAAGACAGATTTGGATGCCCT 3960
OY 4024 GGAGAGACACATGAAGAAATGATTTGCTGAGTACCTGTAACCTGAATCTGAAGT 4083
Db 3961 GGAGAGACACATGAAGAAATGATTTGCTGAGTACCTGTAACCTGAATCTGAAGT 4020
OY 4084 GATGAGAAATGGGATTAATCTTCAAGTGGGGAGACGAGCTTGTGATAGTGAAGC 4143
Db 4021 GATGAGAAATGGGATTAATCTTCAAGTGGGGAGACGAGCTTGTGATAGTGAAGC 4080
OY 4144 CTTGCTCCGCCACTGTAAGATTTCTGATTTTATGATGAAGGACAGCTGCCATGACACAGA 4203
Db 4081 CTTGCTCCGCCACTGTAAGATTTCTGATTTTATGATGAAGGACAGCTGCCATGACACAGA 4140
OY 4204 GACAGACTTATTTGATTCAGAGACCATCCGAGAGCAATTTGACAGCTGACATGCTGAC 4263
Db 4141 GACAGACTTATTTGATTCAGAGACCATCCGAGAGCAATTTGACAGCTGACATGCTGAC 4200
OY 4264 CATTTGCCATGCGCTGACACAGGTTCTAGGCTCCGATAGATTAATGTTGCTGGCCCAAGG 4323
Db 4201 CATTTGCCATGCGCTGACACAGGTTCTAGGCTCCGATAGATTAATGTTGCTGGCCCAAGG 4260
OY 4324 ACAGTGGGAGTTTGAACACCCCATCGGCTCTGCTGACAGCAGATGCCGATTTCTA 4383
Db 4261 ACAGTGGGAGTTTGAACACCCCATCGGCTCTGCTGACAGCAGATGCCGATTTCTA 4320
OY 4384 TGCCATGTTTGTCTGTCAGAGAACAGGTCGCTGTCAAGGGCTGACTCCCTGTGA 4443
Db 4321 TGCCATGTTTGTCTGTCAGAGAACAGGTCGCTGTCAAGGGCTGACTCCCTGTGA 4380
OY 4444 CGAAGTCTTTTCTTTTGTAGAGCATTTGCCATTTCCCTGGGGGGGCCCTTCATCGCG 4503
Db 4381 CGAAGTCTTTTCTTTTGTAGAGCATTTGCCATTTCCCTGGGGGGGCCCTTCATCGCG 4439
OY 4504 TCCCTCAACCAACCTTCCCTTCCATTTATCTTCCGACAGACAGTCCGATTTG 4563
Db 4440 TCCCTCAACCAACCTTCCCTTCCATTTATCTTTCGACAGCAGATTCGGATTTG 4499
OY 4564 CTGTGTGTTTCACTTTTGAAGGAGAGTATATTTGATTTATTTATTTTCCATATTC 4623
Db 4500 CTGTGTGTTTCACTTTTGAAGGAGAGTATATTTGATTTATTTATTTATTTCCATATTC 4559
OY 4624 TGTAAACAAATTTTATTTTGTCTTAATTCACACTTAAAGGTTTCAGAGAACCTTAT 4683
Db 4560 TGTAAACAAATTTTATTTTGTCTTAATTCACACTTAAAGGTTTCAGAGAACCTTAT 4619
OY 4684 TATTAATTTGATGAGGCTTATTAATGAAGCTTTATACGTTAGCTATATCTATATAT 4743
Db 4620 TATTAATTTGATGAGGCTTATTAATGAAGCTTTATACGTTAGCTATATCTATATAT 4679
OY 4744 TCTGTACATAGCCCTATATTTTACAGTGAAGTGAAGCTGTTTTATTTTATTTTAAATAAG 4803
Db 4680 TCTGTACATAGCCCTATATTTTACAGTGAAGTGAAGCTGTTTTATTTTATTTTAAATAAG 4739
|||||

OY	4804	CACGTGCTCTAAAAAA	4819
Dn	4740	CACGTGCTCTAATAACA	4755
RESULT	7		
ABO19003			
LOCUS	ABO19003	5400 bp	mRNA
DEFINITION	Mus musculus MRP5 mRNA, complete cds.		ROD
ACCESSION	ABO19003		01-MAR-2000
VERSION	ABO19003.1	GI:4587084	
KEYWORDS	MRP5.		
SOURCE	Mus musculus cDNA to mRNA.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
TITLE	1 (sites)		
	Suzuki,T., Sasaki,H., Kuh,H.J., Agui,M., Tatsunai,Y., Tanabe,S.,		
	Terada,M., Saijo,N. and Nishio,K.		
	detailed structural analysis on both human MRP5 and mouse mrp5		
JOURNAL	transcripts		
MEDLINE	Gene 242 (1-2), 167-173 (2000)		
REFERENCE	20184734		
AUTHORS	2 (bases 1 to 5400)		
TITLE	Suzuki,T., Kuh,H. and Nishio,K.		
JOURNAL	Direct Submission		
	Submitted (20-OCT-1998) to the DDBJ/EMBL/GenBank databases.		
	Roshino Suzuki, National Cancer Center Research Institute,		
	Pharmacology Division, Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045,		
	Japan (E-mail:csuzuki@gen2.res.ncc.go.jp,		
	Tel.:81-3-3542-2511(ex.4451), Fax:81-3-3542-1886)		
FEATURES	Location/Qualifiers		

```

, polyA_site      5400      /note="18 a nucleotides"
BASE COUNT      1311 a      1318 c      1412 g      1359 t
ORIGIN

```

Query Match		77.9%	Score 3774	DB 94	Length 5400	
Best Local Similarity		88.0%	Pred. No. 0			
Matches 4203		Conservative	0	Mismatches 555	Indels 16	Gaps 8
Qy	1	GGCATCAGCTCTGGGAGAGCGCTGGTTGAGCGGCTGGGCGGTTGCTCTGGAGCAGGGGCGAG	60			
Db	39	GTCGGATGTTCGCGAGCGCTGCCGGGCGTGTTCCTGCTCTGGGGAGGAGCTGGGGTGCAG	98			
Qy	61	GAATTCTGATGTGAACCTACACAGCTGTGAGCCCTCGGAAACCTCCACCTCAGGAAGATCAA	120			
Db	99	AATTTCTGAAGTGAAGGAGGAGTGTGCA--CCCTGGTCTCTGTGTGTGAAGATGAA	155			
Qy	121	GGATTGACATATAGGAAGAAGATATATCATCCCAAGTCTGGGTATAGAGTGTGAGGA	180			
Db	156	AGATATTACATGGGAAAGAAATATATATATATCCCAAGCCCTGGGTACAGAAGTACAGGGA	215			
Qy	181	GAGAACGACACTTCTGGGACGACACAGAGACCGTGAGATTCCAGTTCCAGGACAATCG	240			
Db	216	CAGAGCCCTGTACAGGCGCAACACAGAGACCCCGAGAACCCAGTTCGCGAAGAACAG	275			
Qy	241	AACGTTGAATGGCCAGAAATGCTTGGAACAGACAGCCGAGCCGAGGCGCTCTCTTGA	300			
Db	276	ATCGTTGGAATGCCAAGATGCTCTCTGGAACAGCAGCCCGAGTTGAGGGGCTTTCCCTGGA	335			
Qy	301	TGCCCTCATGCAATTCCTCAGATCAGAACTCCTGGATGAGAGCATCCCAAGGAGAAATACCA	360			
Db	336	TATCTCTGCAATTCCTCATCTCAAAATTCGGACAGAGAGCATTTAAGGGAATAATACCA	395			
Qy	361	TCATGGCTGATGCTCTGGAAGCCCATCCGACATCTTTGGAACACGACGACCCGATGGA	420			
Db	396	CCATGTTAATGTCTCGAAGCCCTTCGCGACACTACCAAGCACAGCACCCAGCTGGA	455			
Qy	421	CAATGCTGGGCTTTTCTCTGTATGACTTTTTCGAGCTTTCTTCTCTGGCCCGTGTGGC	480			
Db	456	CAATGCTGGACTTTTCTCTCATATGACCTTTTCAATGGCTCTCTCTCTGGCCGAGTGGT	515			
Qy	481	CCACAGAAGGGGGAGCTCTCAITGGAAGCGTGTGTCTCTGTCTCAAGCACAGATCTTC	540			
Db	516	TCACAGAAGGGGGAGCTGTTAATGAGAGATGTGTGGCTTGTGTCCAAAGTAAAGATCTTC	575			
Qy	541	TGAGCTGAACCTCAACAAGACTGAGAGACTGTGGCAAGAAGAGCTGAATGAATGGGGCC	600			
Db	576	TGATGTGAACCTCAACAAGACTGAGAGACTGTGGCAAGAAGAGCTGAATGAATGGGGCC	635			
Qy	601	AGACGCTCTCTCCCTGCGAAGGGTTGTGTGATCTTCTGCGACACAGGCTCATCTGTC	660			
Db	636	AGACGCTCTCTCCCTGCGAAGGGTTGTGTGATCTTCTGCGACACAGGCTCATCTGTC	695			
Qy	661	CATGCTGTGCTGATGATACCGCACTGCTGGGCTTCAGTGGACCAAGGCTTCATGGTGAA	720			
Db	696	CATGCTGTGCTGATGATACCGCACTGCTGGGCTTCAGTGGACCAAGGCTTCATGGTGAA	755			
Qy	721	ACACCTTTTGGAGTATACCAAGGACAAGAGCTATACCTGCAGTACAGCTTGTGTAGT	780			
Db	756	ACACCTTTTGGAGTATACCAAGGACAAGAGCTATACCTGCAGTACAGCTTGTGTAGT	815			
Qy	781	GCTGGGCTCTCTGACGAGAAATCTGCGGCTTGTGTGCTTGTGCATGACTTGGGCATT	840			
Db	816	GCTGGGCTCTCTCTGACGAGAAATCTGCGGCTTGTGTGCTTGTGCATGACTTGGGCATT	875			
Qy	841	GAATTAACGAACCGGTGTGCTGCGGGGGGCAATCCAAACATGAGGCAATTAAGAAAGAT	900			
Db	876	GAATTAACGAACCGGTGTGCTGCGGGGGGCAATCCAAACATGAGGCAATTAAGAAAGAT	935			
Qy	901	CTTTAAGTTAAAGACATTTAAAGAGAAATCCCTGGGTAAGCTCATCAATTTGCTCCAA	960			
Db	936	CTTTAAGTTAAAGACATTTAAAGAGAAATCCCTGGGTAAGCTCATCAATTTGCTCCAA	995			
Qy	961	CGATGGGCAAGAAATGTTGAGGACGACAGCCGTTGGACGCTGCGGCTGGGAGGAGACCGT	1020			
Db	996	CGATGGGCAAGAAATGTTGAGGACGACAGCCGTTGGACGCTGCGGCTGGGAGGAGACCTGT	1055			

OY	1021	TGTTGCCATCTTAAAGCATGATTTATATATGTAATTTATTTCTGGAGCCACAAGGCTTCTTG6G	1080
Db	1056	TGTTGCCATCTTGGGACATGATTTATATATGTAATTCATTCCTAGAACCCAGGGCTTCGT6G	1115
OY	1081	ATCAGCTGTTTTTATCTCTTTTACCACAGATGATGTTTGATTCAGGGCTCACAGATA	1140
Db	1116	ATCAGCGGTTTTTATCTCTTTTATATCACAAGATGATGTTGCTGACGGCTACCTGCATA	1175
OY	1141	TTTTAGAGAAAAATGCGTGGCCGCCACAGGATAGCGTCCGCAAGAGATGATGAAGTTCT	1200
Db	1176	TTTTAGAGAAAAAGCGTGTACTGCTCCACAATGACGTGTCCGCAAGAGATGAATGAAGTTCT	1235
OY	1201	TACTTACATTAAATTTATCAAAATGATGCTGGGTCAAGACATTTTCTCAGAGTGTTCA	1260
Db	1236	TACCTACATTAAATTTATTAATGATGATCGTGGGTCAAGCGTTTCTCAGTGTGCA	1295
OY	1261	GAAATATCCGAGAGAGAGAGCGCTGGGATATTGGAAAAAGCCGGGTACTTCCAGAGCATAC	1320
Db	1296	AAAAATCCGAGAGAGAGAGAGCTGGGATATTGGAAAAAGCCGGGTACTTCCAGAGCATAC	1355
OY	1321	TGTGGTGTGGGCTCCCACTTGTGGTGGATATGCGACCGGTGGAGCATCTCGTGGTTCAAT	1380
Db	1356	TGTTGGAGTGGCTCTCTATTGTGGTAACTGATCGCCAGTGTGGTGAAGCTTCTCCGTTCCAT	1415
OY	1381	GACCCGTGGGCTTGATCTGCACACAGCACAGGCTTTCACAGTGGTGAAGCTTCAATTTC	1440
Db	1416	GACCCGTGGGCTTGATCTGCATCTGCGGACACAGGCTTTCACAGTGGTGAAGCTTCAATTTC	1475
OY	1441	CATGACTTTTGTCTTGAAGTAAACCGGTTTTCAAGTAAAGTCCCTCTCAGAGCCTCACT	1500
Db	1476	CATGACTTTTGTCTTGAAGTAAACACACATTCACAGTGAAGTCCCTCTCTGAAAGCATCACT	1535
OY	1501	GGCGTTGGACAGATTTTAAGGTTGTTTTCATATGGAAGAGGTTCCATATGAATAAAGAACAA	1560
Db	1536	GGCGTTGGACAGATTTTAAGGTTGTTTTCATATGAAGAGGTTCCATATGAATAAAGAACAA	1595
OY	1561	ACGAGCGATCTCACATCATCAAGATTAAGATGAATAATGGCACCTTGCGATGGGACCTCTC	1620
Db	1596	ACGCGCCAGTCTCACATCAAGATTAAGATGAATAATGGCACCTTGCGATGGGACCTCTC	1655
OY	1621	CCACTCCAGTATCCAGAACTCGCCCAAGCTGACCCGCCAAAATGAAAAAAGACAAAGAGGC	1680
Db	1656	CCACTCCAGTATACAGAACTCGCCCAAGCTGACCCGCCAAAATGAAAAAAGACAAAGAGGC	1715
OY	1681	TTCCAGGGGCGAAGAAAGAGAAGGTGAGGCACTGCAGCGCACTGAGCATCAGCGGCTCT	1740
Db	1716	TACCAGGGGCGAAGAAAGAGAAGTGCAGGAGCACTGCACACACTGAGCACACAGGCGCTCT	1775
OY	1741	GGCAGAGACAGAAAGGCCACTCTCTCCGTGCAAGTGAAGGAGGCGGCCCACTCCGAAAGAGA	1800
Db	1776	GGCAGAAACAGAAAGGACACTCTCTCTGGAACAGCGACGAGCGGCCCACTCCGAAAGAGA	1835
OY	1801	AGAGAGGAAGACATCTCCACTGGGCCACTCGGCTTACAGAGACACTGCACAGCATCGA	1860
Db	1836	AGAAAGGCAAGAGATCCACACAGAGAGGCGTGGCGCTGCAGAGAGACACTGTACAAATTTGA	1895
OY	1861	TCCTGAGATCCAGAGAGGTTAACTGCTGTGGAATCTGCGGCACTGTGGGAAGTGGAAAAAC	1920
Db	1896	CTTGAATAATGAAGAGGCGCAAACTGGTTGGAACTCTCGGCACTGTGGGAAGTGGAAAAAC	1955
OY	1921	CTCTCTCATTTTACGCCATTTTAAAGGCCAGATGACGTTCTAGAGGGCAGCATTTGCAATAG	1980
Db	1956	CTCTCTCGTTTACGCCATTTTAAAGGCCAGATGACGTTTGGAGGGCAGCATTTGCGCTAG	2015
OY	1981	TGGAACCTTGCTTATGTGGCCAGCAGAGGCTGGATCTCAATGCTACTGTAGAGACAA	2040
Db	2016	TGGGACCTTTTCTATTGTGGCCCAACAGGCGTGTGATTTCTCATTTGCCACTCTGAGAGACAA	2075
OY	2041	CATCCCTGTTTGGGAAGGATATGATGAAGAAAGATCAACTCTGTGCTGAACAGCTGCTG	2100
Db	2076	CATTCCTTTTGGGAAGGATTTGATGAAGAGATATCAACTGATGCTGAATAGCTGCTG	2135

QY	2101	CTGTAGGCGCTGACCTGGCCATTTCTTCCACAGACGCACTGACGGAGATTGGAGACGCAAG	2160
Db	2136	CTGTAGGCGCTGACCTGGCCATTTCTTCCACAGACGCACTGACGGAGATTGGAGACGCAAG	2195
QY	2151	AGCCAACTGAGGCGGTGGAGGCGCAAGAGATCAAGCCTTGGCCGGGCGCTGTATAGGA	2220
Db	2196	AGCCAACTGAGGCGGTGGAGGCGCAAGAGATCAAGCCTTGGCCGGGCGCTGTATAGGA	2255
QY	2221	CAGGAGCATCTACATCTCTGGAGCAACCCCTCAGTGGCTTAGATGCCATATGGGCAATCA	2280
Db	2236	TAGAAGCATCTACATCTCTGGAGTACCCCTCAGTGGCTTAGATGCCATATGGGCAATCA	2315
QY	2281	CATCTTCATATAGTCATATCCGGAACATCTCAAGTCCAAAGACGTTCTGTTTATACCA	2340
Db	2316	CATCTTCATATAGTCATATCCGGAAGCGCTCAAGTCTTAGAGCGTTCTGTTTATACCA	2375
QY	2341	CCATTTACAGTACCTGGTTACGTGTATGAAGATGATCTTCAATGAAGAAGGCGGTATATAC	2400
Db	2376	CCAGTTACAGTATCTGGTTACGATTTGTATGAGTGATCTTCAATGAAGAAGGCGGTATATAC	2435
QY	2401	GGAAGAAGGCGCCCATATGAGAACATGATGAATTAATTAATGCTACTATGCTACCATTTTAA	2460
Db	2436	AGAGAGAGTACCCATATGAGAGGCTGATGAACATTAATTTGGGATTTACGATTTTAA	2495
QY	2461	TAACTGTGTGGGAGAGACACCGCCAGTTGAGATCAATTAAGAAAAGAAAACGATGG	2520
Db	2496	TAACTGTGTGGGAGAGACACCCCGATGAGATTAATTTGAAAAAGAAAGTACATGG	2555
QY	2521	TTTACAGAAAGATCACAAACAAAGGCTCTTAACAGGATCAATTAAGAAAGAAAGAAAGC	2580
Db	2556	TTTAC - - AAAATACAAACAAAGGCGCTTAAGCCAGGCTAGTAAAGAAAGAAAGAAAGC	2612
QY	2581	AGTAAAGCCAGAGAAAGGAGAGCTTGTCAGCTGGAAGAAAGGAGGAGGTTCTAGTCC	2640
Db	2613	GCTGAAGTCGAGAAAGGAGGAGCTTGTCAGAGGTGGAAGAAAGGAGGAGGTTCTAGTCC	2672
QY	2641	CTGGTCAATATATGTTGTCTACATCCAGGCTGCTGGGGGCGCCCTTGCCATTTCTGGTAT	2700
Db	2673	TTGGTCAATATGTTGTGTCTACATCCAGGCTGCAAGGGGCGCCCTTGCTGCTGTAT	2732
QY	2701	TATGGCCCTTTTCACTGCTGAATGTAGGACAGACCGCCTTACAGCAGCTGCTGTATGTTA	2760
Db	2733	CATGGTCTCTTCACTGCTGAATGTGGGACAGACATGCTTACAGCAGCTGCTGTATGTTA	2792
QY	2761	CTGGATCAAGCAAGGAAGCGGAAACCACTGATCTCAAGGGAACAGACCTGCTGATG	2820
Db	2793	CTGGATCAAGCAAGGAAGCGGAAACAGCAAGTATCAAGGGAACAGACAGCTTGTATG	2855
QY	2821	TGACAGCATGGAAGACATCTCATATATGCAGTACTATGTCACAGCATCTACGCGCTTCCAT	2880
Db	2853	TGACAGCATGGAAGACAAACCCCTTCAATGCAAGTACTATGTCACAGCATCTACGCGCTTCCAT	2912
QY	2881	GCGACTATGCTGATCTGAAAGCCATGAGAGATTGCTTTGTCAAGGCAAGCTCGC	2940
Db	2913	GCGACTATGCTGATCTTGAAGCCATGAGAGATTGCTTTGTCAAGGCAAGCTCGC	2972
QY	2941	AGCTTCCCTCCCGGCTGATACAGAGCTTTTCCGAAGATCTTCAAGACCTCAATGAAGTT	3000
Db	2973	AGCTTCCCTCCCGGCTCATATATGAGCTATTCCGAAGATCTTGAAGACCTCAATGAAGTT	3032
QY	3001	TTTTTACACAGACCCCAAGAGGAGATTTCAACAGGTTTTTCCAAAGACATGATGAAGT	3060
Db	3033	TTTTTATCTACCCCAACAGGAAGATTTCAACAGGTTTTTCCAAAGACATGATGAAGT	3092
QY	3061	TCAGCTGCGGCTGCGCTTCCAGGCGCAGATGTTCACTCCAAACGTTATCTGTGTTCTT	3120
Db	3093	GGATGTCGCGCTGCGCTTCCAGGCTGAGATGTTATTCAGAAATGTAAATCTGTGTTCTT	3152
QY	3121	CTGTGTGGGATATGTCAGAGAGCTTCCCGTGTCTCTTGTGGCAGTGGGGCGCCCTGT	3180
Db	3153	CTGTGTGGGATATGTCAGAGCTTCCCGATCTTCCATGTTCTCTGTGGCGGTGGGCGCTCTCT	3212
QY	3181	CATCTCTTTTCACTCTCTGCAACATTTGTCTCAAGGCTCGATTCGGAGCTTACAGCTCT	3240

[illegible]

Db	4293	GACATTGCCATGGCCGTGCACACAGTTCTGGGCTCTGACAGGATCATGGTGCAGGCCA	4352
QY	4321	GGGACAGTGGTGGAGTTTGACACCCCATGGTCTCTGTCCACAGCAGTCCGAT	4380
Db	4353	GGGACAGTGGTGGAGTTTGACACCCCATGGTCTCTGTCTGATATGATTCAGAT	4412
QY	4381	CTATGACATGTTTCTGTCGACAGAACAGAGTGGCTGTCAAGGAGCTGACCTCCCTGT	4440
Db	4413	CTATGACATGTTTCTGTCGACAGAACAGAGTGGCTGTCAAGGAGCTGATCTCCCTG	4471
QY	4441	TGACGAGTCTCTTTCTTTAGACATATGCGATTCCTGCTGGGGGGGCCCTTCATC	4500
Db	4472	CCCTGAGATGTC--TTCCTGAGACATATGCGATTCCTGCTGGGGGGGCCCTTCATC	4525
QY	4501	GCGTCTCTCTACCGAACCTTGGCTTCTGATTTATCTTTGCGACAGAGTCCGAT	4560
Db	4526	ATCCGTCCTGCTCTAAACCTTGGCTTCCAGTTTATCTCTCAACACCATTCAGAT	4585
QY	4561	TGGCTGTGTTTCTCATTATTTAGGAGAGCATATTTGATTTATTTATTCATAT	4620
Db	4586	TAGC-TGTTGTTTCTCATTATTTAGGAGAGCATATTTGATTTATTTATTCATAT	4644
QY	4621	TCATGTAAACAAATTTAGTTTGTCTTATTTGCACTC-TAAAGGTTTCAGGGAAACG	4679
Db	4645	TCATGTAAATGAATTTAGTTTGTCTTATTTGCACTCTTAAAGGTTTCAGGGAAACG	4704
QY	4680	TTATTTATTAATG-TATCAGAGGCTTATTAAGAGCTTTATACGTATGCTATAT	4738
Db	4705	TTACTATAAAGTTTATCAGTGGCTTATAAAGAGCTTTATATGTTAGCTATATAT	4764
QY	4739	ATAATCTGATACAGCTATATTTTACAGTAA	4772
Db	4765	ATAATCTGATACAGCTATATTTTACAGTAAA	4798
RESULT	8		
AB005659			
LOCUS	AB005659	4939 bp	mRNA
DEFINITION	Homo sapiens SMRP mRNA, complete cds.		PRI
ACCESSION	AB005659		
VERSION	AB005659.1	GI:2554609	
KEYWORDS	SMRP.		
SOURCE	Homo sapiens bone marrow CDNA to mRNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 4939)		
JOURNAL	Suzuki,T.		
REFERENCE	Submitted		
TITLE	Submitted (10-JUL-1997) to the DDBJ/EMBL/Genbank databases.		
JOURNAL	Toshitiro Suzuki, National Cancer Center Research Institute,		
REFERENCE	Pharmacology Division; Tsukiji 5-1-1, Chuo-ku, Tokyo 104, Japan		
AUTHORS	(E-mail:tsuzukikgan2.res.ncc.go.jp, Tel.03-3542-2511,		
TITLE	Fax:03-3542-1886)		
JOURNAL	2 (sites)		
REFERENCE	Suzuki,T., Nishio,K., Sasaki,H., Kurokawa,H., Saito-Ohara,F.,		
AUTHORS	Ileuchi,T., Tanabe,S., Terada,M. and Saijo,N.		
TITLE	cDNA cloning of a short type of multidrug resistance protein		
JOURNAL	homologue, SMRP, from a human lung cancer cell line		
REFERENCE	biochem. Biophys. Res. Commun. 238 (3), 790-794 (1997)		
MEDLINE	97472289		
FEATURES			
SOURCE	Location/Qualifiers		
	1..4939		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="3"		
	/map="3q27"		
	/tissue_type="bone marrow"		
gene	737..3577		
	/gene="SMRP"		
CDS	737..3577		
	/note="a short type of multidrug resistance protein"		

```

homologue"
/codon_start=1
/protein_id="BAA22887.1"
/db_xref="GI:2534610"
/translation="MKNAITLAMPSSSHSSSIONSPKLTTPKMKDKKRNKSKKEKRYOLR
TERQAVLAIEQKHLILDSDERPSEEEKHILGHILRLORTLSIDLEIOEGVLGI
CGSVGSKTSLISIIIGOMTLLEGISIAISGFPAVAQAOALINTLNDLIFGEYVE
ERYNSVNSGCLRPDLILPSSDLTEIGEGANSGORISIAIRLADRSYIID
DPLSALDAHVNHIINSATIRKHLKSTVLYTHOLOYLVDDEYIKKECITRGTH
EELMNLNGDYATIFNNLLLEPEVEINSKETSQSKSQDKPKTSVKEKAAP
EEQOLVLEEGGGSVPWSYGYVIOAGDPLAEIVLMALFMLNVGSTAFSTWMLSYM
IKGGSGNTVTTRNETSVSDSMKDNPMQYASISYALMALFMLNVGSTAFSTWMLSYM
RASRLHDELFRILRSPKPDPTPTGRILNRSKMDVDVRLPQAEFIONVL
VPCVGMIAGVPEMPLVAVGPVILVPSVLIVSVRLIRLKRDNITQSPLSITSS
IOGLATHATNKGOEFLHRYOELDDNOAFPEFTCAMRLAVLIDLITSLITTLG
MYLMHGOIPPAITAGLAISTAVOLITGLFTVRLASTETAEATNVSERINHYIKTSLI
APARINKRASPMPPOGEVTFENAEKMYRENLPLVAKVSTIKPEKIGIVRGTS
GKSILGMLRLVELSGGCIKIDVIRISDGLADLRKSLIIPQEPVLFSGTVSND
PFIQYTEDOIMDALERTHMEKLAQLPLKLESEYEMNGDNFVSEROLLIARLLRH
CKTILIDEATAAMDTEFDLIIETIREAPADCTMLNLAHRLHYLSGDRIMVLAQGV
VERDTPSVLLSNSRFPYAFMAENKVAIKG"
4939
polyA_site
/note="15 A nucleotides"

```

BASE COUNT 1174 a 1224 c 1249 g 1292 t
 ORIGIN

Query Match 76.48; Score 3703.8; DB 85; Length 4939;
 Best Local Similarity 96.98; Pred. No. 0;
 Matches 3844; Conservative 0; Mismatches 7; Indels 116; Gaps 2;

```

OY 968 CAGGAATGTTGAGGCGACGCGTGGCAGCCGTCGGCTGGAGGACCCGTTGTC 1027
    |||||||
Db 1 CAGGAATGTTGAGGCGACGCGTGGCAGCCGTCGGCTGGAGGACCCGTTGTC 60
OY 1028 ATCTAGGCATGATTTAATATGTAATATTCCTGGACCAAGAGCGTCTGGATCAGCT 1087
    |||||||
Db 61 ATCTAGGCATGATTTAATATGTAATATTCCTGGACCAAGAGCGTCTGGATCAGCT 120
OY 1088 GTTTTATACCTCTTTACCCAGCAATGATGTTTGCATCAGCGCTCAGCAATATTCAGG 1147
    |||||||
Db 121 GTTTTATACCTCTTTACCCAGCAATGATGTTTGCATCAGCGCTCAGCAATATTCAGG 180
OY 1148 AGAAAATCGTGCGCCGACGAGATGATGTCGCAAGATGAAATGATTTACTTAC 1207
    |||||||
Db 181 AGAAAATCGTGCGCCGACGAGATGATGTCGCAAGATGAAATGATTTACTTAC 240
OY 1208 ATTTAAATTTATCAAAATGTATGCTGGTCAAAAGCATTTTCTCAGAGTGTTCAGAAAATC 1267
    |||||||
Db 241 ATTTAAATTTATCAAAATGTATGCTGGTCAAAAGCATTTTCTCAGAGTGTTCAGAAAATC 300
OY 1268 CGCGAGGAGGAGCGTGGATTTGAAAAAGCGGGTACTTCCAGAGATCAGTCTGGGT 1327
    |||||||
Db 301 CGCGAGGAGGAGCGTGGATTTGAAAAAGCGGGTACTTCCAGAGATCAGTCTGGGT 360
OY 1328 GTGGCTCCCATTTGTTGTTGATGTCGAGCGGTGACCTTCTGTCATATGACCGTG 1387
    |||||||
Db 361 GTGGCTCCCATTTGTTGTTGATGTCGAGCGGTGACCTTCTGTCATATGACCGTG 420
OY 1388 GCGTTTCGATCTGACAGCAGCAGGCTTTCACAGTGTGACAGTCTTCAATTCATGACT 1447
    |||||||
Db 421 GCGTTTCGATCTGACAGCAGCAGGCTTTCACAGTGTGACAGTCTTCAATTCATGACT 480
OY 1448 TTTGCTTTGAAAGTAAACCGCTTTTCACTAAAGTCCCTCTCAGAAAGCCTCAGTGGCTGT 1507
    |||||||
Db 481 TTTGCTTTGAAAGTAAACCGCTTTTCACTAAAGTCCCTCTCAGAAAGCCTCAGTGGCTGT 540
OY 1508 GACGAGATTTA----- 1517
    |||||||
Db 541 GACGAGATTTAAGCTTCTTCCACCTGTATAGCCTGATGTTGTTTCTTACGTAGTGTACGC 600
OY 1518 ----- 1517

```

```

Db 601 CCTAGGCTTGTCTCCCTGCTGTGATCCAGTGTCTCATCTTTGCAACCTCTCTCACAACTC 660
OY 1518 -----AGAGTTTGTTCCTAATGGAAGAGGCTTCACATGTATAAAGAACAAACGACCTCC 1572
    |||||||
Db 661 CTATCAGAGTTTGTTCCTAATGGAAGAGGCTTCACATGTATAAAGAACAAACGACCTCC 720
OY 1573 TCACATCAAGTATGAGATGAAAAAATGCCACCTTGGCATAGGGACCTCTCCACATCAGTAT 1632
    |||||||
Db 721 TCACATCAAGTATGAGATGAAAAAATGCCACCTTGGCATAGGGACCTCTCCACATCAGTAT 780
OY 1633 CCAGAACTCGCCCAAGCTGACCCCAATGAAAAAAGACAAGAGGCTTCACAGGGGCAA 1692
    |||||||
Db 781 CCAGAACTCGCCCAAGCTGACCCCAATGAAAAAAGACAAGAGGCTTCACAGGGGCAA 840
OY 1693 GAAAGAGAAGTGAAGGCGAGCTGACGCGCAGCTGAGCATAGCGGGTGTGGCAGAGCAAA 1752
    |||||||
Db 841 GAAAGAGAAGTGAAGGCGAGCTGACGCGCAGCTGAGCATAGCGGGTGTGGCAGAGCAAA 900
OY 1753 AGGCCACTCTCTCGGACAGTGAAGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1812
    |||||||
Db 901 AGGCCACTCTCTCGGACAGTGAAGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
OY 1813 CATCCACTGAGGCGCACTGCGCTTACAGAGGACACTGCAAGCATGATCTGAGATCCA 1872
    |||||||
Db 961 CATCCACTGAGGCGCACTGCGCTTACAGAGGACACTGCAAGCATGATCTGAGATCCA 1020
OY 1873 AGAGGTTAACTGGTTGGAATCTGCGGCACTGTGGGAAAGTGGAAAAACCTCTCATTTTC 1932
    |||||||
Db 1021 AGAGGTTAACTGGTTGGAATCTGCGGCACTGTGGGAAAGTGGAAAAACCTCTCATTTTC 1080
OY 1933 AGCCATTTAGGCGCAGATGACGCTCTGAGAGGCGAGATTCGAATGCAATGAGAACCTTCCG 1992
    |||||||
Db 1081 AGCCATTTAGGCGCAGATGACGCTCTGAGAGGCGAGATTCGAATGCAATGAGAACCTTCCG 1140
OY 1993 TTATGTGCGCCAGCAGGCGCTGGATCTCAATGCTACTGTGAGAGCAACATCTGTTGG 2052
    |||||||
Db 1141 TTATGTGCGCCAGCAGGCGCTGGATCTCAATGCTACTGTGAGAGCAACATCTGTTGG 1200
OY 2053 GAAGGAATATGATGAAGAAAGATTAAGTCTGTGTGAAGAGCTGCTCCAGAGCGCTGGA 2112
    |||||||
Db 1201 GAAGGAATATGATGAAGAAAGATTAAGTCTGTGTGAAGAGCTGCTCCAGAGCGCTGGA 1260
OY 2113 CCTGSCAATCTCTCCACAGAGGACCTGACGAGATTTGAGAGCGAGAGGCAACCTGAG 2172
    |||||||
Db 1261 CCTGSCAATCTCTCCACAGAGGACCTGACGAGATTTGAGAGCGAGAGGCAACCTGAG 1320
OY 2173 CGGTGGGAGCGCCAGAGATCAGCTTGCCTGGGCTTGTATGATGAAGAGCATCTA 2232
    |||||||
Db 1321 CGGTGGGAGCGCCAGAGATCAGCTTGCCTGGGCTTGTATGATGAAGAGCATCTA 1380
OY 2233 CATCTGAGACGACCCCTCAGTGTGCTTATGATGCCATGTGGGCAACCATTTCAATAG 2292
    |||||||
Db 1381 CATCTGAGACGACCCCTCAGTGTGCTTATGATGCCATGTGGGCAACCATTTCAATAG 1440
OY 2293 TGCCTCCGGAACATCTCAAGTCCAAAGCAAGTCTGTTTGTACCAACATTTACAGA 2352
    |||||||
Db 1441 TGCCTCCGGAACATCTCAAGTCCAAAGCAAGTCTGTTTGTACCAACATTTACAGA 1500
OY 2353 CCTGTTACTGTGTATGAAGTATCTTCAATGAAGAGGCTGTATTAAGGAAAGGAGC 2412
    |||||||
Db 1501 CCTGTTACTGTGTATGAAGTATCTTCAATGAAGAGGCTGTATTAAGGAAAGGAGC 1560
OY 2413 CCATGAGAACTGATGAATTTAATGTTGACATATGCTTACCATTTTAAATACCTTGTCT 2472
    |||||||
Db 1561 CCATGAGAACTGATGAATTTAATGTTGACATATGCTTACCATTTTAAATACCTTGTCT 1620
OY 2473 GGGAGAGACCGGCGAGTGTGATCAATTTCAAAAAAGAAACCACTGGTTCACAGAGAA 2532
    |||||||
Db 1621 GGGAGAGACCGGCGAGTGTGATCAATTTCAAAAAAGAAACCACTGGTTCACAGAGAA 1680
OY 2533 GTACAAAGACAAAGGCTCTTAAACAGATCAATAAAGAGAAAAAGCAATTAAGCCAGA 2592
    |||||||
Db 1681 GTACAAAGACAAAGGCTCTTAAACAGATCAATAAAGAGAAAAAGCAATTAAGCCAGA 1740

```

```
OY 2593 GGAAGGCGAGCTGTGTCAGCTGGAAGAAAAGGCGAGGGTTCACTGCCCTGGTCAGTATA 2652
|||||
Db 1741 GGAAGGGGAGCTGTGTCAGCTGGAAGAAAAGGCGAGGGTTCACTGCCCTGGTCAGTATA 1800
|||||
OY 2653 TGGTGTCTACATCCAGGCTGTGTGGGGGCCCTTGGCAATTCCTGGTTATATGAGCCCTTTT 2712
|||||
Db 1801 TGGTGTCTACATCCAGGCTGTGTGGGGGCCCTTGGCAATTCCTGGTTATATGAGCCCTTTT 1860
|||||
OY 2713 CATGCTAATGTATGAGCAGCAACCGCTTCAGCACTGGTGTGTGAGTTACTGATCAAGCA 2772
|||||
Db 1861 CATGCTAATGTATGAGCAGCAACCGCTTCAGCACTGGTGTGTGAGTTACTGATCAAGCA 1920
|||||
OY 2773 AGGAAGCGGGAAACCACTGTGACTGAGGGAGAGAACCTCGGTGAGTGCACAGATGAA 2832
|||||
Db 1921 AGGAAGCGGGAAACCACTGTGACTGAGGGAGAGAACCTCGGTGAGTGCACAGATGAA 1980
|||||
OY 2833 GGAATATCCATATGAGATATATGCGAGATATGAGCCCTCCATGAGAGATGATGCT 2882
|||||
Db 1981 GGAATATCCATATGAGATATATGCGAGATATGAGCCCTCCATGAGAGATGATGCT 2040
|||||
OY 2893 GATCCTAAGAACCACTGTGAGAGATGTCTTTGTCAAGGGCAGCTGCGAGCTTCCTCCG 2952
|||||
Db 2041 GATCCTAAGAACCACTGTGAGAGATGTCTTTGTCAAGGGCAGCTGCGAGCTTCCTCCG 2100
|||||
OY 2953 GCTGCAATGACAGCTTTTCCGAGAGATCCTTGAAGCCCTATGAAATTTTTCACACAGAC 3012
|||||
Db 2101 GCTGCAATGACAGCTTTTCCGAGAGATCCTTGAAGCCCTATGAAATTTTTCACACAGAC 2160
|||||
OY 3013 CCCACAGGGGAGATTCGAAACAGATTTTCCAAAGACATGATGATGAGTGCAGCGGCT 3072
|||||
Db 2161 CCCACAGGGGAGATTCGAAACAGATTTTCCAAAGACATGATGATGAGTGCAGCGGCT 2220
|||||
OY 3073 GCGCTTCAGGCGCAGATGTTTCATCAGAAAGTTATCCTGGTGTCTCTCTGTGTGGGAT 3132
|||||
Db 2221 GCGCTTCAGGCGCAGATGTTTCATCAGAAAGTTATCCTGGTGTCTCTGTGTGGGAT 2280
|||||
OY 3133 GATGCGAGAGATCTTCCCGTGGTTCCTTGTGCGAGTGGGGCCCTTGTCACTCTTTTC 3192
|||||
Db 2281 GATGCGAGAGATCTTCCCGTGGTTCCTTGTGCGAGTGGGGCCCTTGTCACTCTTTTC 2340
|||||
OY 3193 AGTCTGACATGTCGTCGAGGGTCTGATTCGGGAGCTGGAAGGTGTCGCAATATTCAC 3252
|||||
Db 2341 AGTCTGACATGTCGTCGAGGGTCTGATTCGGGAGCTGGAAGGTGTCGCAATATTCAC 2400
|||||
OY 3253 GCACTACACCTTCTCTCCACATCAGCTCCAGCATACAGAGGCCCTTGCCACATCCAGGC 3312
|||||
Db 2401 GCACTACACCTTCTCTCCACATCAGCTCCAGCATACAGAGGCCCTTGCCACATCCAGGC 2460
|||||
OY 3313 CTACAAATAAAGGAGAGATTTCTGCACAGATACCAGAGCTGCTGGATGCAACCAAGC 3372
|||||
Db 2461 CTACAAATAAAGGAGAGATTTCTGCACAGATACCAGAGCTGCTGGATGCAACCAAGC 2520
|||||
OY 3373 TCCCTTTTGTGTATAGTGTGATGCGGTGCTGGGCTGGACCTCATCAG 3432
|||||
Db 2521 TCCCTTTTGTGTATAGTGTGCGATGCGGTGCTGGGCTGGACCTCATCAG 2580
|||||
OY 3433 CATGCGCCTATCACCAACAGGGGCTGATGATGTTCTTATGCAAGGCGAGATTCGCC 3492
|||||
Db 2581 CATGCGCCTATCACCAACAGGGGCTGATGATGTTCTTATGCAAGGCGAGATTCGCC 2640
|||||
OY 3493 AGCCTATGCGGGTCTGCGCATCTTATGCTGTCCAGTTAAACGGGGCTGTCCAGTTTAC 3552
|||||
Db 2641 AGCCTATGCGGGTCTGCGCATCTTATGCTGTCCAGTTAAACGGGGCTGTCCAGTTTAC 2700
|||||
OY 3553 GGTGAGACTGGCATCTGAGACAGAAAGCTGATTCACCTCGGTGAGAGAGATCAATFACATA 3612
|||||
Db 2701 GGTGAGACTGGCATCTGAGACAGAAAGCTGATTCACCTCGGTGAGAGAGATCAATFACATA 2760
|||||
OY 3613 CATTAAGACTGTGCTTGAAGACCTGCGCAAGATTAAAGAAACAGGCTCCCTCCCTGA 3672
|||||
Db 2761 CATTAAGACTGTGCTTGAAGACCTGCGCAAGATTAAAGAAACAGGCTCCCTCCCTGA 2820
|||||

OY 3673 CTGCCCCAGGAGGAGAGATGACCTTTGAGAACGCAAGATGAGTACCGAGAAAACCT 3732
|||||
Db 2821 CTGCCCCAGGAGGAGAGATGAGTACCTTTGAGAACGCAAGATGAGTACCGAGAAAACCT 2880
|||||
OY 3733 CCTCTGCTCTTAAGAAAGATATCCTTCACGATCAAACTTAAGAGAAAGTGTGATTTGT 3792
|||||
Db 2881 CCTCTGCTCTTAAGAAAGATATCCTTCACGATCAAACTTAAGAGAAAGTGTGATTTGT 2940
|||||
OY 3793 GGGGCGACAGATCAGGAAAGTCCCTGCTGGGGATGGGCCCTTTCGTCGTGAGATT 3852
|||||
Db 2941 GGGGCGACAGATCAGGAAAGTCCCTGCTGGGGATGGGCCCTTTCGTCGTGAGATT 3000
|||||
OY 3853 ATCTGAGAGCTGCAATCAAGATGATGAGTGAAGATATGATGATTTTGGCCTTGGCAGCT 3912
|||||
Db 3001 ATCTGAGAGCTGCAATCAAGATGATGAGTGAAGATATGATGATTTTGGCCTTGGCAGCT 3060
|||||
OY 3913 CCGAAGCAAACTCTCTATCATCTCTCAAGAGCCGCTCTGTCAGTGGCACTGTCAATC 3972
|||||
Db 3061 CCGAAGCAAACTCTCTATCATCTCTCAAGAGCCGCTGTGTCAGTGGCACTGTCAATC 3120
|||||
OY 3973 AAATTTGGACCCCTTCAACCACTGTAAGAGACAGATTTGGAGTCCCTGAGAGGAC 4032
|||||
Db 3121 AAATTTGGACCCCTTCAACCACTGTAAGAGACAGATTTGGAGTCCCTGAGAGGAC 3180
|||||
OY 4033 ACACATGAAGAAATGTATGTCTCAGCTACCTCTGAACCTGTAATCTGAAGTGAAGAA 4092
|||||
Db 3181 ACACATGAAGAAATGTATGTCTCAGCTACCTCTGAACCTGTAATCTGAAGTGAAGAA 3240
|||||
OY 4093 TGGGGATAACTTCTCATGTGGGGAGAACGCACTCTGTGCTAGACTAGAGCCCTGCTCCG 4152
|||||
Db 3241 TGGGGATAACTTCTCATGTGGGGAGAACGCACTCTGTGCTAGACTAGAGCCCTGCTCCG 3300
|||||
OY 4153 CCACTGTAAATTTCTGATTTTATGTAAGAACCCACAGCTGCCATGACACAGACAGACTT 4212
|||||
Db 3301 CCACTGTAAATTTCTGATTTTATGTAAGAACCCACAGCTGCCATGACACAGACAGACTT 3360
|||||
OY 4213 ATGATTTCAAGACCAATCCGAGAAAGATTTGCAAGCTGTAACCTGTCAGCAATTTGCCA 4272
|||||
Db 3361 ATGATTTCAAGACCAATCCGAGAAAGATTTGCAAGCTGTAACCTGTCAGCAATTTGCCA 3420
|||||
OY 4273 TCGCCTGACACAGGTTTCTAGGCTCCGATAGGATATGTTGCTGGGCCAGAGGACGTGT 4332
|||||
Db 3421 TCGCCTGACACAGGTTTCTAGGCTCCGATAGGATATGTTGCTGGGCCAGAGGACGTGT 3480
|||||
OY 4333 GGAATTTGACACCCCATGCTGTCTGTCTCCACAGACAGTTCCGATTTCTATGCAATGTT 4392
|||||
Db 3481 GGAATTTGACACCCCATGCTGTCTGTCTCCACAGACAGTTCCGATTTCTATGCAATGTT 3540
|||||
OY 4393 TGTGCTGCAAGAAACAAGTGCCTGTCGAAGGGCTGACTCCCTGTTGCGAAGTCTC 4452
|||||
Db 3541 TGTGCTGCAAGAAACAAGTGCCTGTCGAAGGGCTGACTCCCTGTTGCGAAGTCTC 3600
|||||
OY 4453 TTTTCTTTAGACATTTGCAATCCCTGCTGAGGGGCCCTTCAATCGGCTCTCTAC 4512
|||||
Db 3601 TTTTCTTTAGACATTTGCAATCCCTGCTGAGGGGCCCTTCAATCGGCTCTCTAC 3659
|||||
OY 4513 CGAAACCTTGCTTTTCGATTTTATCTTTGCGACAGCAGTTCCGGAATGGCTGTGTGT 4572
|||||
Db 3660 CGAAACCTTGCTTTTCGATTTTATCTTTGCGACAGCAGTTCCGGAATGGCTGTGTGT 3719
|||||
OY 4573 TTTCACTTTAGGAGAGCATATTTGATTTATTTATTTATTCATATTCATGTAACAA 4632
|||||
Db 3720 TTTCACTTTAGGAGAGCATATTTGATTTATTTATTTATTTATTCATATTCATGTAACAA 3779
|||||
OY 4633 AATTTAGTTTGTCTTAAATTTGCACTCTAAAGGTTCCAGGAAACGTTATATATTTGT 4692
|||||
Db 3780 AATTTAGTTTGTCTTAAATTTGCACTCTAAAGGTTCCAGGAAACGTTATATATTTGT 3839
|||||
OY 4693 ATCAGAGGCTTAATATGAGCTTTATACGTGTACCTATATCTATATATTTCTGTACAT 4752
|||||
Db 3840 ATCAGAGGCTTAATATGAGCTTTATACGTGTACCTATATCTATATATTTCTGTACAT 3899
|||||
OY 4753 AGCCTATATTTACAGTAAATATGTAAGCTGTATTTATTTATTTAAATTAAGCACTGTGCT 4812
|||||
```



```
Db 931 CCGAACCTAAAGACATTAAAGAAAAGTCCCTGGGGAGCTTATCAATATCTCTCCAA 990
QY 961 CGATGGGCAGAGAAATTGTTAGGCAGACGCCGTTGGACGCTCTGCTGGAGACCCGT 1020
Db 991 CGATGGGCAGAGAAATTGTTAGGCAGACGCCGTTGGAGAGCTCTGCTGGAGACCCGT 1050
QY 1021 TGTGGCATCTGGGCGATGATTTATATGTAATTTCTGGGACCAACAGGCTTCGCGG 1080
Db 1051 TGTGGCATCTGGGCGATGATTTATATGTAATTTCTGGGACCAACAGGCTTCGCGG 1110
QY 1081 ATGAGCTTTTATCCCTCTTTTACCGAGAAATGTTTGCATACAGGCTCACAGATA 1140
Db 1111 ATGGCGCTCTTATCCTCTTTTATCCGGCAATGATGTTGTGCACAGGCTAACCGGTA 1170
QY 1141 TTTTCAGAGAAAATGCGTGGCCGCGACGATGAACGTGTCCAGAAATGAATGAATGCT 1200
Db 1171 TTTTCAGAGAAAATGCTGTACCGCCACAGACGACGCTGTCCAGAAATGAATGAATGCT 1230
QY 1201 TACTTCAATTAAATTTTCAAAAATGATGCTGGGTCAAGCAATTTTCTCAGAGTTCA 1260
Db 1231 TACTTCAATTAAATTTTCAAAAATGATGCTGGGTCAAGCAATTTTCTCAGAGTTCA 1290
QY 1261 GAAATTCGCGAGAGAGAGCGTGGATTTGAAAAAGCCGGTACTTCCAGACATCAC 1320
Db 1291 AAAAATCCGAGAGAGAGAGCGTGGATCTCGAGAAAGCCGGTACTTTCAGACATCAC 1350
QY 1321 TGTGGTGTGGCTCCCATTTGTGTGTGATTTGCCAGCGGTGTGACCTTCTCTTCAAT 1380
Db 1351 CGTGTGAGTGTGCTCCCATTTGTGTGTGATTTGCCAGCGGTGTGACCTTCTCCCAT 1410
QY 1381 GACCCGTGGCTGTGATGTGACAGACAGACAGCGTTTTCACAGTGTGTACAGTTCATTC 1440
Db 1411 GACCCGTGGCTGTGACAGAGCGGCGAGCGCTTTCACAGTGTGTACAGTTCATTC 1470
QY 1441 CAGTCTTTTGTGTTAAATTAACCCGTTTTCAGTAAAGTCCCTTCAGACAGCTCAGT 1500
Db 1471 CAGTCTTTTGTGTTAAATTAACCCATTTCTCAGTAAAGTCCCTTCAGACATTCCT 1530
QY 1501 GCGTGTGACAGATTTAAGAGTTGTTTCAATGGAAGAGGTTTCATGATTAAGAACAA 1560
Db 1531 TCGCTGTGACAGATTTAAGAGTTGTTTCAATGGAAGAGGTTTCATGATTAAGAACAA 1590
QY 1561 ACCAGCCAGTCTCACATCAAGATAGAGATGAAAAATGCCACCTTGGCATTTGGACTCTC 1620
Db 1591 ACCGCGAGTCTCACATCAAGATAGAGATGAAAAATGCCACCTTGGCATTTGGACTCTC 1650
QY 1621 CCACTCCAGTATCCAGAACTGCGCCCAAGCTGACCCCAAAATGAAAAAGACAGAGGCG 1680
Db 1651 CCACTCCAGTATCCAGAACTGCGCCCAAGCTGACCCCAAAATGAAAAAGACAGAGGCG 1710
QY 1681 TTTCCAGGGGCAAGAAAGAGAGTGAAGGACGCTGACGCGCATGAGCATCAGATCAGGCGT 1740
Db 1711 TTTCCAGGGGCAAGAAAGAGAGTGAAGGACGCTGACGCGCATGAGCATCAGATCAGGCGT 1770
QY 1741 GGGAGAGCAAAAGGGCCACCTCTCTGAGACAGTGAAGAGAGCGGCGCATGAGAGAGGA 1800
Db 1771 GGGAGAGCAAAAGAGACCTCTCTCTGAGACAGTGAAGAGAGCGGCGCATGAGAGAGGA 1830
QY 1801 AGAAGGCAAGCACATCCACCTTGGGCACTGCGCTTACAGAGACATGACAGATCGA 1860
Db 1831 AGAAGGCAAGCACATCCACGAGGAGCATGCGTCTGACAGAGACATGACAGATTTGA 1890
QY 1861 TCTGGAGATCCAAAGAGGTAATCTGGTTGGAATCTGCGGCAAGTGTGGAAAGTGAAGAAC 1920
Db 1891 CTTAGAGATTTGAAGAGGCAAACTGTTGGAATCTGCGGCAAGTGTGGAAAGTGAAGAAC 1950
QY 1921 CTCTCTCATTTTCAGCCCTTTTAGGCGCAGATGAGCGCTTCTAGAGGGCAGCATTTGCAATCAG 1980
Db 1951 CTCTCTCATTTTCAGCCCTTTTAGGCGCAGATGAGCGCTTTCGAGGGCAGCATTTGCGCTCAG 2010
QY 1981 TGGAACTTGTGTTATGCGCCAGCAGACGAGCTGGATCTCAATGCTACTCTAGAGAGCAA 2040
Db 2011 TGGGACTTTTGTCTAGCTGCGCCACAGAGGCGCTGATTTCTCAAGCCACGCTGAGAGAGCAA 2070

QY 2041 CATCTCTTTTGGGAAGGAATTTGATGAAGAAAGATACAACTCTCTGTGAAACAGCTCTG 2100
Db 2071 CATCTCTTTTGGGAAGGAATTTGATGAAGAAAGATACAACTCTCTGTGAAACAGCTCTG 2130
QY 2101 CCTGAGGCTGACCTGGGCATCTTTCCAGAGAGGACCTGACGGAATTTGGAGAGCGAGG 2160
Db 2131 CCTGAGGCTGACCTGGGCATCTTTCCAGAGAGGACCTGACCTGACGGAATTTGGAGAGCGAGG 2190
QY 2161 AGCCAACTGAGCGGTGGGAGCGCCAGAGAGATCAGCTTGGCCGCTTTGATTAAGTA 2220
Db 2191 AGCCAACTGAGGTGGGAGCGCCAGCGAATCAGCTTGGCCGCTTTGATTAAGTA 2250
QY 2221 CAGAGCATCTACATCTGTGAGAGACCCCTCAGTGCCTTAGATGCCATGTGGGCAACCA 2280
Db 2251 CAGAGCATCTATATCTTGGATGATCCCTCAGTGCCTTAGATGCCATGTGGGCAACCA 2310
QY 2281 CATCTCAATAGTCTATTCGCGAAGATCTCAATCCAGACATTTCTGTTTGTACCCA 2340
Db 2311 CATCTTCAACAGTCTATTCGCGAAGCGTCTCAAGTCTAAGAGGTCCTGTTTGTACACA 2370
QY 2341 CCAATTACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
Db 2371 CCAATTACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2430
QY 2401 GGAAGAGGACCCATGAGAGACTGATGATTTAATGTGATCTATCTTACATTTTAA 2460
Db 2431 AGAAGAGGATCCACAGAGGAGTGTGATGATTTAAGTGAAGTCACTTACATTTTAA 2490
QY 2461 TAACTGTGCTGGAGAGAGACACCGAGTTGAGATCAATTTCAAAAAGGAAACCACTGG 2520
Db 2491 TAACTGTGCTGGAGAGAGACACCGAGTTGAGATCAATTTCAAAAAGGAGTAACTGG 2550
QY 2521 TTTCAAGAGAAAGTCAACAGAACAAAGGCTCTTAAAAAGAGATCAATTAAGAAAGAAAAAGC 2580
Db 2551 TTTCAAC---AAAAATCAACAGAACAAAGGCGCTTAAGCCAGAGGTCATGAAAGAAAGAGGCG 2607
QY 2581 AGTAAAGCCAGAGAGAGGAGGAGCTTGTGACGCTGGAAGAGAAAGGAGGTTGATGCC 2640
Db 2608 AGTAAAGCCAGAGAGAGGAGGAGGAGCTGTGACGCTGGAAGAGAAAGGAGGAGGCTGCC 2667
QY 2641 CTGTCAGTATATGTTGCTTCAATCCAGGCGCTGTGGGGGCCCTTGGCATTTCTGTTAT 2700
Db 2668 TTGTCAGTATATGTTGCTTCAATCCAGGCGCTGTGGGGGCCCTTGGCATTTCTGTTAT 2727
QY 2701 TATGGCCCTTTTCAATGCTGATGATGAGGAGAGCAGCCCTTCAAGCACTGTGTTGAGTTA 2760
Db 2728 CATGTCCTCTTCAATGCTGATGATGAGGAGAGCAGCTTCAAGCACTGTGTTGAGTTA 2787
QY 2761 CTGATCAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2820
Db 2788 CTGATCAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2847
QY 2821 TGACGATGAGAGAGCAATCCATATGACAGTACTATGCGACAGCTTACGCGCTCTCCAT 2880
Db 2848 TGACGATGAGAGAGCAATCCCTTCTGACAGTACTATGCGACAGCTTCTCTCCAT 2907
QY 2881 GGCAGTATGCTGATCTCTGAAAGGCAATTCGAGAGTGTCTTTGTCAAGGGCAGCTGCG 2940
Db 2908 GGCAGTATGCTGATCTCTGAAAGGCAATTCGAGAGTGTCTTTGTCAAGGGTACACAG 2967
QY 2941 AGCTTCTCTCCGCTGATGAGAGGCTTTTCCGAAGGATCTTCCGAAGGCTTATGAAGT 3000
Db 2968 AGCTTCTCTCCGCTGATGAGAGGCTTTTCCGAAGGATCTTCCGAAGGCTTATGAAGT 3027
QY 3001 TTTTTCAGAGCCGCCCAAGGAGGATTTCAAGAGGTTTCCAAAGACATGATGAAGT 3060
Db 3028 TTTTTCATCTACCCCAAGAGAGGATTTCTACAGAGTCTTTTCCAAAGACATGATGAAGT 3087
QY 3061 TGACGTGCGGCTGCCCTTCAAGGCGAGATGTTCAATCAGAAAGTTATCTGCTGCTTCT 3120
Db 3088 GGAAGCTGCTGCTGCAATTCAGAGGCTGAGATGTTATTCAGAAATGATGATCTGCTGCTT 3147
```


QY	3121	CTGTGGGAAATATCGACAGAAATCTTCCCGTGGTTCCTTGGGACAGTGGGGCCCTTGT	3180
Db	3148	CTGTGGTGGAAATATCGCAGGGGTCTTCCCATGTTCTCTTGTGGCAGTGGGGCCCTTCT	3207
QY	3181	CATCCTCTTTCAGTCCGTGCACATTGTCTCCAGGGTCCCGATTCGGGAGCTGAAGCCTCT	3240
Db	3208	CATCCTCTTCTCAGTTCTTACACTTGTCTCCAGGGTCCCGATTCGGGAGCTGAAGCGCTT	3267
QY	3241	GGACAAATATCAGCGAGTCACACTTTCCTCTCCACATCATCGTCCAGCATACAGGGCCCTTGC	3300
Db	3268	GGACAAATATCAGCAAAATCTCTTCCCTCCCATATTCACGTCCACATTCACAGGGCCCTGCG	3327
QY	3301	CACCATCCAGGCTACAAATTAAGGGGAGAGATTCTCGACACATACAGAGAGCTGTGGA	3360
Db	3328	CACCATCCATGCTCTTATPACAAAAGGAGAGATTCTTACACAGTATACAGAGACTTGTGGA	3387
QY	3361	TGACAAACCAAGCTCTCTTTTCTTTTGTATTCAGTGTGAGATGCGGTGGCTGGCTGTGCGCT	3420
Db	3388	TGACAAACCAAGCCCCCTTCTTCTTGTTCACCTGTGCAATGAGGTGGCTGGCAGTGGCT	3447
QY	3481	GCAGATTCCCCACGCTATGACGGGTGTGGCCATCTCTTATGCTGTCCAGTTTACGGGGCT	3540
Db	3508	GCAGATTCCTCTCAGCGTACGGGGCCCTCCGCATCTCTACGCTGTGCAGTTTACCGGGACT	3567
QY	3541	GTTCACATTACGGTGCAGACTGGCATCTGACAGACAGAACTGCATTCACTCGGTGGGAGAG	3600
Db	3568	CTTTCACATTTACGTTAGACTGTGCGGAGACAGAGACAGCGTTTCACCTCAGTGGAGAG	3627
QY	3601	GATCAATCCTACTACATTAAAGCTCTGTCTTGAAGACACTGGCAGAAATTAAGAACAGGC	3660
Db	3628	GATCAACACTACTACTAAAGCTCTGTCTTGAAGACACTGGCAGAAATTAAGAACAGGC	3687
QY	3661	TCCCTCCCTGACTGCGCCCGACGAGGAGAGAGTGACCTTTTGAAGACGACAGATGAGSTA	3720
Db	3688	TCCCTCCCATAGACTGTGCGCCCGACGAGGAGAAATTAACCTTTTGAAGATGGCGAAATGAGTA	3747
QY	3721	CCGGAAGAAACCTCCCTGTGCTCTTAAGAAAGATTCCTACAGATCAAAACCTTAAGAA	3780
Db	3748	CCGGAAGAAATCTCCCTGTGCTCTTAAGAAAGTGTCTTACACATCAAGCCCAAGGAA	3807
QY	3781	GATTGGCATTTGTGGGGGCGACAGATCAGAGGAAGTCTTCGTGGGGATAGGCCCTTTCG	3840
Db	3808	GATTGGCATTTGTGGGAAGAACAGGTCAGAGGAAGTCTTCTTAAGGATAGGCCCTTTCG	3867
QY	3841	TCGTGTGAGCTTATCTGAGGCTGCATCAAGATTGATGAGGATGAGTAATCACTGATATTTGG	3900
Db	3868	CTGTGTGAGATTATCTGAGGCTGCATCAAGATTCATGATGATGATGATGATGATGATTTGG	3927
QY	3901	CTTTGGCGACTCTCGAAGCAAACTCTCTATTCATCTCTCAAGAGCCGGGTGTCTTCAATGG	3960
Db	3928	CTTTGGCGACTCTCGAAGCAAACTGCATCATCTCTCGAGGCGCAGTGTCTTCAAGGG	3987
QY	3961	CACGTGCAGATCAAAATTTTGGAGCCCTTCAACAGATCATCTGTAAGAACAGATTTGGATGTC	4020
Db	3988	CACAGTGCAGATCAAAACTGSAACCCCTTCAACAGATCATCAAGAAAGCAGATCTGGGATGTC	4047
QY	4021	CTTGGAGAGACACACATGAAAGATATGTCTCAGCTACCTCTGAAACTTGAATCTGA	4080
Db	4048	CTTGAAGAGAACACACATGAAAGATATGTATGCTCAGCTACCTCTGAAACTTGAATCTGA	4107
QY	4081	AGTATGAGAAATGSGGATATCTTCTCAGTGGGGAGACGGCCTTGTGTGATAGCTAG	4140
Db	4108	AGTATGAGAAACGGGACAACTTCTCTGTGGGGAACGGCAGTTGTATGATAGCAAG	4167
QY	4141	AGCCCTCTCTCGGCACGTATAGATTTGTGATTTAGATGAAGCACAGCTGCATGGACAC	4200
Db	4168	GGCCCTCTCTAGCTACGTATAGATTTGTGATTTAGATGAAGCTACAGCCGCTATGAGAC	4227
QY	4201	AGAGACAGACTTATTTGATTCAGAGACCATCGGAGAGCACTTTGGACACTGTACCATGCT	4260

Db	4228	AGAGCAGACTTACTGATCCAAAGACCATCCGGGAAACATTTGCTAGCTGCACATGCT	4287
QY	4261	GACCATGCCCCATCGCCCTGCACACAGGTTCTAGGCTCCGATAGAGATTATGTCGCGCCA	4320
Db	4288	GACCATGCCCCATCGCCCTGCACACAGGTTCTGCGCTGTATAGATCATGTCGCGCCA	4347
QY	4321	GGGACAGGTGGTGGAGGTTTGACACCCCATCGTCTTCTGTCCAAACGACAGTTCGGATT	4380
Db	4348	GGGACAGGTGGTGGAGGTTTGACACCCCATCGTCTTCTGTCCAAACGACAGTTCGAATT	4407
QY	4381	CTATCCCATGTTTGGTGGTGGAGAGAAAGGTGCTGTCAAGGGCTGACATCCCTCGT	4440
Db	4408	CTATCCCATGTTTGGTGGTGGAGAGAAAGGTGCTGTCAAGGGCTGACATCTTCC---	4464
QY	4441	TGACGAGTCTCTTTCTTTTAGAGCATTTGCCATTCCTGCTGGGGCGGGCCCTTCATC	4500
Db	4465	ACGCTGGAAGTCTCTTCTTGGAGGCTTGCCATTCCTGCTGGTGGGGCCCTCATCTC-	4523
QY	4501	GCGTCTCCCTACCGAACCCTGCTTCTTCGATTTTATCTTTGCGACACAGTTCGGAT	4560
Db	4524	-----GTCTGCTGAAACCTCGCTCTTCTGCTT--TCTCTCAGCAGTCAAGGGT	4576
QY	4561	TGGCTTGTGTGTTTACCTTTTAGGAGAGTCATATTTGATATGATATATTCATAT	4620
Db	4577	TAGC-TGTGTGCTCAGCTTTTAAAGAAATCATATTTGATATTTATATTCCTAT	4635
QY	4621	TCATCTAAACAAATTTTGGTTTCTGCTTAATTCAGCTC-TAAAGGTTGAGGACCG	4679
Db	4636	TCATCTAAATGAATTTTGGTTTCTGCTTAATTCAGCTCCTTAAGAGTTGAGGACCA	4695
QY	4680	TTATTTAATATGTTATCAGAGGCTTAATGAAGCTTTATACGTAGCTATATATATATA	4739
Db	4696	TTACTATATA-CGTATCAGGCGCTTAATGAAGCTTTATACGTAGCTATATATATA	4754
QY	4740	TATATCTGACTATACCTATATTTACAGTGAAGATGAAGCTTTATTTATATATAA	4799
Db	4755	TATATCTGACTATACCTATATTTACAGT--AAATGTAAGTGGTTATTTCTATTTAAAG	4812
QY	4800	TAAAGCATGTGCTAAAAA 4818	
Db	4813	CGAGCATGTGCTAAACACA 4831	
RESULT	10		
LOCUS	A64695	1761 bp	DNA
DEFINITION	Sequence 61 from Patent W09731111.		
ACCESSION	A64695		
VERSION	A64695.1	GI:4530760	
KEYWORDS			
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1761)		
AUTHORS	Oude,E.R., Paulsuma,C.C., Bosma,P.J., Borst,P., Evers,R., Kool and Marcel.		
TITLE	A FAMILY OF ORGANIC ANION TRANSPORTERS, NUCLEIC ACIDS ENCODING THEM, CELLS COMPRISING THEM AND METHODS FOR USING THEM		
JOURNAL	INTROGENE BV (NL)		
COMMENT	other publication AU 1736697 19970910.		
FEATURES	Location/Qualifiers		
SOURCE	1..1761		
BASE COUNT	427 a 448 c 469 g 413 t		4 others
ORIGIN	/organism="unidentified" /db_xref="taxon:32644"		
Query Match	34.5%;	Score 1674.4;	DB 9; Length 1761;
Best Local Similarity	99.5%;	Fred. No. 0;	
Matches 1722; Conservative	0;	Mismatches	1; Indels 7; Gaps 4;

QY 1969 CATTCGATCAGTGGAACTTCGCTTATGTGGCCAGCAGGCTGATCTCAATGCTAC 2028
 DB 1 CATTCGATCAGTGGAACTTCGCTTATGTGGCCAGCAGG - CTGGATCTCAATGCTAC 59
 QY 2029 TCTAGAGACACATCTCTGTTGGGAAGAAATATGATGAAGAAAGATCAACTCTGTGCT 2088
 DB 60 TCTAGAGACACATCTCTGTTGGGAAGAAATATGATGAAGAAAGATCAACTCTGTGCT 119
 QY 2089 GAACAGTGTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2148
 DB 120 GAACAGTGTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 179
 QY 2149 TGGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2208
 DB 180 TGGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 239
 QY 2209 CTGTATATGATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2268
 DB 240 CTGTATATGATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 299
 QY 2269 TGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2328
 DB 300 TGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 359
 QY 2329 GTTT - GTTACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2385
 DB 360 GTTTNNNGTTACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 419
 QY 2386 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2445
 DB 420 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 479
 QY 2446 TGTACCATTTTATATACCTGTTGCTGGAGAGACACCCGAGTGTGATCAATTCATA 2505
 DB 480 TGTACCATTTTATATACCTGTTGCTGGAGAGACACCCGAGTGTGATCAATTCATA 539
 QY 2506 AAAGAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2565
 DB 540 AAAGAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 599
 QY 2566 AAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 2625
 DB 600 AAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 659
 QY 2626 GCAGGAGTTCAGTGGCTGAGTGGTATATGATGATGATGATGATGATGATGATGATGATG 2685
 DB 660 GCAGGAGTTCAGTGGCTGAGTGGTATATGATGATGATGATGATGATGATGATGATGATG 719
 QY 2686 GGCA - TTCTGCTGATATATAGGCTTTTCATG - CTGAATGTAGGAGGAGGAGGAGGAGG 2742
 DB 720 GGCAATTCCTGCTGATATATAGGCTTTTCATGAGGCTGATGAGGAGGAGGAGGAGGAGG 779
 QY 2743 CACCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2802
 DB 780 CACCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 839
 QY 2803 GAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2862
 DB 840 GAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 899
 QY 2863 CATTCAGGCTTCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2922
 DB 900 CATTCAGGCTTCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 959
 QY 2923 TGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2982
 DB 960 TGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1019
 QY 2983 TCGAAGGCTTATGAAGTTTATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3042
 DB 1020 TCGAAGGCTTATGAAGTTTATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1079

QY 3043 CAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3102
 DB 1080 CAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1139
 QY 3103 CGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3152
 DB 1140 CGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1199
 QY 3163 GCGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3222
 DB 1200 GCGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1259
 QY 3223 TCGGAGGCTGAAGGCTGGAACAATATGACGAGTACCTTCTCTCCATCAGGCTG 3282
 DB 1260 TCGGAGGCTGAAGGCTGGAACAATATGACGAGTACCTTCTCTCCATCAGGCTG 1319
 QY 3283 CAGCATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3342
 DB 1320 CAGCATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1379
 QY 3343 ATACAGGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3402
 DB 1380 ATACAGGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1439
 QY 3403 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3462
 DB 1440 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1499
 QY 3463 GATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3522
 DB 1500 GATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1559
 QY 3523 TGTCCAGTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3582
 DB 1560 TGTCCAGTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1619
 QY 3583 ATTCACCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3642
 DB 1620 ATTCACCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1679
 QY 3643 CAGAAATTAAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3692
 DB 1680 CAGAAATTAAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1729

RESULT 11
 LOCUS A64696 2167 bp DNA PAT 29-MAR-1999
 DEFINITION Sequence 62 from Patent WO9731111.
 ACCESSION A64696
 VERSION A64696.1 GI:4530761
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 2167)
 AUTHORS Oude, E.R., Paulusma, C.C., Bosma, P.J., Borst, P., Evers, R., Kool and Marcel.
 TITLE A FAMILY OF ORGANIC ANION TRANSPORTERS, NUCLEIC ACIDS ENCODING THEM, CELLS COMPRISING THEM AND METHODS FOR USING THEM
 JOURNAL INTRIGENE BV (NL)
 COMMENT Other publication AU 1736697 19970910.
 FEATURES
 source 1..2167
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 494 a 511 c 511 g 626 t 25 others
 ORIGIN
 Query Match 21.7%; Score 1054.2; DB 9; Length 2167;
 Best Local Similarity 96.0%; Pred. No. 7.9e-239;

Matches	2108; Conservative	0;	Mismatches	1895; Indels	125; Gaps
QY	284	GAGGCGCTCTCTTGATGCTCCATGCATCTCTCAGCTCAGATCCTGGATGAGAGCAT	343		
Db	217	GATGGCCCCCTGGAGTCCAGCAAGAGAAATCCTGAGGCTCCAGGGAGGGACACTCTGCCA	276		
QY	344	CCCAAGGGAAGTACCATCATGCGCTTGAGTGTGTCGAAGCCCATCG-----GACTACT	397		
Db	277	CCGTGGGGGAAGTATGATGCTGCTTGGAAACATGATATTCCTTCCGTCGCCAAGCCGAGG	336		
QY	398	TGCAAAACACAGCACCCAGTGGACATCTGGGGCTTTTTCCTGTATGACTTTTTCGTGG	457		
Db	337	TTTCTGCCCCCAGCCCCCTGGACATCTGGCCCTGTTCTCTTACCCTCAGCCGTGCATGG	396		
QY	458	CTTCTCTCTGCGCCCGTGGCCCAACAAGAGGGGGAGCTTCATATGGAAAG-A-GGTGG	516		
Db	397	CTCAC-----CCCGCTCATGATCCAAAGCTTACGGAGTGGCTTAGATGAGAACCATCC	452		
QY	517	GTCCTGTCCAAAGCAGAGCTTCTTGACGTGAACGCAAGAGACTAGAGAGACTGTGGCA	576		
Db	453	TCCACTGTCACTGCATGATGCTCAGACAAAATGTCCAAAGGCTTCAACGCCCTTGGGA	512		
QY	577	AGAAAGCTGAATGAAGTGTGGGCCAGACGCTGCTCCCTGGCAAGGTTGTGTGATCTT	636		
Db	513	AGAAAGAACTTCAGAGCGAGGAGATTGAAAACCTTCAGTGCTTGCTGATCTGAGGTT	572		
QY	637	CTGCGCGACACAGGCTCATCTGTCCATCGTGGCTGCCTGATGATCAGACAGCTGGTGGCTT	696		
Db	573	CCAGAGAACAAAGTTGATTTTCGATGCACTTGTGGGACATCTGCTTCTGTGATGGCAGTGT	632		
QY	697	CAGTGGACACGCTTCATGCTGTAACACCTCTGTGAGATACCCAGGCAACAGACTTAA	756		
Db	633	ACTGGGCGCAATATGATATTATPACCAAAATCCTGGAAATTTATGAAAGAGACAGTTGGGGA	692		
QY	757	CCTGCAGTACAGCTTGTGTTAGTACTGGGCCCTCCTCTGACGGAATCTGCGGCTTGG	816		
Db	693	TGTTGTCCATGGAGTGGAGACTGCTGTGGCCCTTTTCTCTCCGATGTGAAGCTCT	752		
QY	817	GTCGCTTGCACCTGACTTGGGCAATTGAATACCGAACCGGTGTCCGCTTGGGGGGCCAT	876		
Db	753	GAGTTTCTCCACAGTTGTGATCATCAACCAAGCAGACAGCATCAGATTCCGAGCACCTGT	812		
QY	877	CCTAACATGCGATTTAAGAAAGTACTCTTAAGTTTAAAGAACTTAAAGAGAAATCCCTGGG	936		
Db	813	TTCCCTCTTGGCTTGGAGAAAGCTATCCAAATTTAAGTCTGTAAATACATCACTCAGG	872		
QY	937	TGAGCTCATCAACATTTGCTCTCCACGATGGGCGAGAAATGTTTGAAGCGCAGACCGTTGG	996		
Db	873	AGAGGCACTACAGCTTCTTACCGGAGTGAATAACTGTTGAAGGGGGGTGTGTAATGG	932		
QY	997	CAGCTGCTGGCTGGAGGACCCGTTGTGGCATCTTAGCAGATTTAATATGTAATTAAT	1056		
Db	933	ACCCTTAAGTATGATACCTGCGCATTCGCTGGTCATCTGCACACTTTCTTCTACTTCAT	992		
QY	1057	TCTGGAGCAACAGAGCTTCTGGGATACAGCTGTTTATTAACCTCTTTTACACGACATGAT	1116		
Db	993	TATTGGATACACTGCAATTATTGGCACTGTATGCTATCCCGGTTTTTCCACTGGAGGT	1052		
QY	1117	GTTTGCATCAGCGCTCACAGCAFTATTTCAGAGAAATGCGTGGCCGCCACGGAATGACG	1176		
Db	1053	ATTCTATGACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTGAGGCCACAGG	1112		
QY	1177	TGTCCAGAAGATGATGATGAGTCTTACTACTATTAATTAATTAACAANAATGATGCCGGGT	1236		
Db	1113	CATCGGTGACAGAGTGAAGTTCTCACTTGGCATTTAAGCTGATTTAAATATATACACATGGGA	1172		
QY	1237	CAAAAGCTTTTCTCAGAGTGTTCAGAAAATCCGAGAGAGAGCGTGGATATTGGAAAA	1296		
Db	1173	GAAAACATTTGCAAAAATCATTTGAGAGACTTAAGAAAGAAAGAAAGAACCTATTGGAGAA	1232		
QY	1297	AGCGGGATCTTCCAGAGCATCACTGTGGGTGGCTCCCATTTGTGTGTGATTTGCCAG	1356		
Db	1233	GTGCGGGCTGTGTCAGAGACCTCAGAGTATAACCTTGTTCATCATCCCGCAGTGGCCAC	1292		

QY	1357	CGTGGTACCTTCCTCTGTTTATATATATGACCCCGGGCTTCGATCTGACAGACAGACAGCTTT	1413
Db	1293	AGCGGTCTGGGTTCTCATCTCACACATCTTTAAAGCTGAAACTCACAGGCTCAATGGCTT	1352
QY	1417	CACAGTGTGACAGTCTTCATTTCCATGACTTTTGTTGAAAGTAAACCGCTTTTCAGT	1476
Db	1353	CAGATCTGATGCCCTCTGAAATCTCCCTGGGCTGTCAAGTTCCTTGTCCTATGAC	1412
QY	1477	AAAGTCCCTCTCAGAAAGCCCTAGTGGCTGTGTGACAGATTTAAGAGTTTGTTCATATGA	1538
Db	1413	CAAAAGTCTCAGCAATTCCAAGTCTCAGATGATGAGTTTCAAGAAATTTTCTCCACAGA	1472
QY	1537	AGAGTTTCACATGATTAAGAAACAACAGCAGCAGTCCCTCATCAAGATAGAGATGAAAA	1596
Db	1473	GAGCCCTGTTTCTATGTCTCAGACATTTTAAAGAACCCACCAAGCTGTGCTTTGAGGA	1532
QY	1597	TGCCACCTTGGCATGGGACTCTCCCATCTCCAGTATCCAGAACTGCCCAAGCTGACCC	1656
Db	1533	GGCCACCTTTCATGTGGCAACAGACCTGTCCCGGGATC-----	1569
QY	1657	CAAAATGAAAAAGACAAGAGGCTTCCAGGGGCAAGAAAGAAAGTGGACACTGCA	1716
Db	1570	-----	1563
QY	1717	GCGCATGAGCATCAGCGGTGCTGGCAGACAGAAAGGCCACTCTCTCTGGACATGA	1776
Db	1570	-----GTCAATGGGCACTGGAGCTGGAGAGAAAGGGGACTGTCTGAGGGGATGAC	1622
QY	1777	CGACGGGCCAGTCCCGAABAGGAAGAACGACACATCCACTGGGCGGCACTCGCTT	1836
Db	1623	CAGGCTTAGAGATGCTCCTCGGGCCAGGGAAGAGGAAGAGTCTGGGCCAGAG-----	1677
QY	1837	ACAGAGGACACTGCACAGCATGCTGAGATCCAAAGGGGTAAACTGTGTGAATCTG	1896
Db	1678	-----TTGCACAAGATCAACCTGGGGTGTCCAAAGGGATGATGTTAGGGGTCTG	1727
QY	1897	CGCAGTGTGGGAATGGAAAAACCTCTCATTTAGCATTTTAAAGCCATAGACCT	1956
Db	1728	CGGGAACCGGGGAGTGTAGAGACAGCTGTTGTACGCATCTGGAGAGATCACTT	1787
QY	1957	TCTAGAGGCGAGCTTTCATCTAGTGGAACTTCCGTTATGTGGGCCACAGGCTGTGAT	2016
Db	1788	GTCTGAGGGCTCGGTGGGGGTGCAGGGAACCTGGCTATGTCTCCACAGGCTGGAT	1847
QY	2017	CTCTCAATGCTACTGTGAGAGACAACATCCTGTTTGGGAAGAAATATGATGAAGAAAGTA	2078
Db	1848	CGTCAAGCGGGAACTCAGGGGAAGAACTCTCATGAGGAGGCGCATATGACAAAGCCCGATA	1907
QY	2077	CAACTGTGTGTGAACAGCTGTGCTGAGAGGCTGACCTGGGCATCTTCCAGCAGGGA	2136
Db	1908	CTCTCAGGTGCTCCACTGCTGCTCTCGAATCGGGACTGTGGAACCTTCTCCCTTGGAGA	1967
QY	2137	CCTGACGAGATTTGGAGAGGAGAGGCCAATCTGAGCGGTGGGACAGCCGACAGAGATCAG	2196
Db	1968	CATGACAGAGATTTGGAAGAGGGGGCCCAACTCTCTGGGGGGCAGAAACAAGATCAG	2027
QY	2197	CCTTGGCCCGGGCTTGTATGATGACAGAGACATCTCAATCTGTGAGAGACCCCTCAGC	2255
Db	2028	CTGTGGCCCGGCCCTCTAATTTCCAGCTCGTCAATCTACTGTGTGGAGGACCCCTGTCTGC	2087
QY	2257	CTTAGATGCTCCATGTGGGCAACACATCTTCATTAATGTATCTCCGAACATCTCAAGTC	2318
Db	2088	TGTGTGACGCTCACTGTGGGAAGCAATTTTGGAGAGTGCATTTAAGAAACACTCAGGGG	2144
QY	2317	CAGAAGTTCCTGTTTGTATACCCACAGTTACAGTACTGTGACTGTGATGAAGTAT	2378
Db	2148	GAAACGCTCTCCAGGTGACCCACAGCTGTGACTTGAATTTTGTGGCCAGTCTAT	2207
QY	2377	CTTCATGAAGAGGCGTGTATTACGGAAGAAAGGCCACCAAGAGAACTGATGAATTTAA	2433
Db	2208	TTTGTGAAAAATGGAAAAATCTGTGAAAAATGGAACCTCAAGTGAATTAATCAGAAAA	2267

QY	2437	TGGGACATATGCATACACATTTTTTAATACCTGTGGTGGGAGAGACACCGCCAGTTGAGAT	2496
Db	2268	GGGGAAATATGCCCAACTTATCCAGAGATGCACAAGGAAGCCATTCCGACATGTTGCA	2327
QY	2497	CAATTCAAAAAAGGAACCAAGTGGTTCACAGAAAGATCACAAAGCAAGGGTCCATAAAC	2556
Db	2328	GGACACAGCAAAAGATATGACAGAGAACCAAGGTAGAAAGTCAAGGCTCTGGCACCTCCCT	2387
QY	2357	AGGATCATATAAAGAGAAAAAGCACTAAAGCCAGAGAGAGGGCAGCTTTTGCACTGGA	2616
Db	2388	GGAAGAGTCTCTCCAAAGGAAATGCTGT-----GCCGGAGCATCAGCTCACACAGAGAGA	2441
QY	2617	AGGAAAAGGCGCAGGTTCAAGTGGCCGTGCTAGTATATGTTGTCTCATATCCAGGCTGCTGG	2676
Db	2442	GGAGATGGAAGAAAGGCTCCTTATGTTGGAGGGTCTACCAACAATATCCAGGAGCTGG	2501
QY	2677	GGGCCCCCTTGGCATTCCTGGTTATATATATGAGCCCTTTTCATATGCTGAATATGAGCAGACGC	2736
Db	2502	AGGTTACATGTCCTCTGTCATATATTTCTTTTGTGGTGATGCGTCTCTTAAGCAT	2561
QY	2737	CTTCAGACACTGGTGGTTGAGTTACTGATCAAGCAAGGAGCGGAAACACCACCTGTGAC	2796
Db	2562	CTTCAGCTTCGTGGTGGCTGAGCTACGTGTTGAGAGAGGCGCTCGGAGACCAATATAGACCG	2621
QY	2797	TCGAGGAGAGAGACC---TCGGTGAATACAGCATGMAAGCAATTCCTCATATGACATA	2853
Db	2622	AGAGAGCAATGAACCAATGGCACAACCTGGGCACACATGCGACAAATCCCAACTGTCTCTT	2661
QY	2854	CTATGCGCAGCATCTAGAGCCCTCTCCATATGGCAGTCAATGCTGATCTCGAAAGGCATTCGAG	2913
Db	2682	CTACACACTGTGTAAAGGGGTCAAGCCCTCGCTCCATCTGTGTGTGGGGTCTGTCTCTC	2741
QY	2914	AGTTGTCTTTGTCAAGGGCAGCGTGCAGCTTCCTCCGGCTGCATGACAGACTTTTCCG	2973
Db	2742	AGGATTTTCCACCAAGGTCAAGGAGGCAATCCAGGCGCTCGACACAAAGCTCTTCAA	2801
QY	2974	AAGGATCCTCGAAGGCCCTATGAAAGTTTGTGACAGACAGCCCAAGAGGAGATCTCAA	3033
Db	2802	CMAAGTTTCCGCTGCCCCCAATGATGTTTCTTGAACACCAATCCCAATTAAGGCGGGTTTTGAA	2861
QY	3034	CAGGTTTCCAAAGACATGATGAAATATACGTGGGGCTGCGCTTCCAGGCGAGATGTT	3093
Db	2862	CTGCTTGGCAGGGGACTTGSAAAGCTGGACACAGCTCTTGCCCATCTTTTCAGAGCAATT	2921
QY	3094	CATCCAGAACGTTATCCGTGTCTTCTTCTGTGTGGGAATGATCGAGAGTCTTCCGTG	3153
Db	2922	CTGTGCTCTCTCTTATATGATGATGCCCGCTCGTGTGATGTGACAGTGTCTCTCATTA	2981
QY	3154	GTTCTCTGTGGCAGTGGGGCCCTGTCTCATCTCTTTTAAGTCTGCACATTTGTCTCAG	3213
Db	2982	TATCTCTTAAATGGGAGCCATAATCATATGTTTATTTGCTTCATTATATATAGATTTCAA	3041
QY	3214	GGTCTCATTTGGGAGCTGMAAGGCTGTGACAAATATCAGCAGTCAACCTTCCCTCCCA	3273
Db	3042	GGAGGCGCATGGGTGTGTTCAAGAGACTGGAGAACTATACCGGCTCTCTTATTTTCCCA	3101
QY	3274	CATCAGCTCCAGATACAGGGCTTCCACACATCCACGCGCTCAATTAAGGGCAGAGATT	3333
Db	3102	CATCTCAATTTCTGTGCAAGGCGCTGAGCTTCATCATCTATGTATGGAATAAATGAAAGACT	3161
QY	3334	TCCTCACAGATACCAAGAGCTGCTGGATGACAACAAGCTCTCTTTTGTGTTTGTAGGTG	3393
Db	3162	CATAGCGCAATTAAGAGGCTGACATGATCGGAGATTAATCTACCTGCTGTGTTCTATTC	3221
QY	3394	TGGCATGCGGGGTGGCTGTGGGGTGGAGCCATCATACACATCGCCCTCATATACCACAC	3453
Db	3222	TTCCACACAGATGAGCTGATGAGCTGTGAGATCATGACCAACCTTGATCACTTGGCCGT	3281
QY	3454	GGGCGCATATATATGTTCTTATGACAGGGAGATATCCCCAGCGCTATGGGGCTCTGCCAT	3513
Db	3282	TGCCCTGTGTGCTTTTGTGGCAATTTCTCTACACCCCTTACTCTTTAAAGTATGCGTGT	3341
QY	3514	CTCTTATGCTGTCCAGTTAAAGGGGCTGTTCCAGTTTACGGTCAAGCTGCATCTGAGAC	3573

DB	3342	CAACATGCTGCTGCACTGTGGCCGTCACGCTTCCAGGCGCACTGGCCCGATTGGCTTGGAGCA	3401
QY	3574	AGAAGCTGATTCACCTCGGTGGAGAGATCAATCATTAAGTCTGTCCCTGCA	3633
Db	3402	AGAGGCAAGATTCACGGCTGTAGAGAGATTAATGCAAGTACATGAAGATGTGTCCGA	3461
QY	3634	AGCACTGCCAGATTAAAGAACAGAGCTCCCTCCCTGACTGTCGGCCCGCCAGAGAGAGGT	3693
Db	3462	AGCTCCCTTTACATGGAAAGGCAAAAGTGTCCCGAGGGGTGGCCACAGCATGGGGAAAT	3521
QY	3694	GACCTTTGAGACGACAGATGAGTACCGAGAGAAACCTCCCTGCTCATTAAGTAAAGT	3753
Db	3522	CATATTTTCAGGATTTATCATATGAATATCAGAGACAAACACACCACCGTCTTCACGGAT	3581
QY	3754	ATVCTTCACGATCAAACTTAAAGAGAAATTTGGCATTTGTGGGGCGGACAGATCAGGAA	3813
Db	3582	CAACCTGACCATCGCGGCGCCACGAAATGGTGGGCATCTGGGAAGAGGCGCTCGGAA	3641
QY	3814	GTCTCGCTGGGATAGGCCCTCTTCCGCTGTGGGAGATTATGTGAGGCTGCATCAAT	3873
Db	3642	GTCTCTCTTGGGACATGGCTCTCTTCCGCTGTGGGAGCCCATGAGAGCGGATTCAT	3701
QY	3874	TGATGAGTGAATCATAGTATATTTGGCTTCGCGCACTCCGAAACAAATCTCATCAT	3933
Db	3702	TGACGGGCTGACATTTTGCAGCATTCGCGCTGGAGAGACTTGGCGTCCAAAGCTTCAGTAT	3761
QY	3934	TCTCAAGAGCCGGTGTCTTCAAGTGGACATGTCAATCAAAATTTGACCCCTTCACCA	3993
Db	3762	CCCTCAAGATTCAGTGTCTCTCAGAGAACATCAATCAATCAATCAATCAATCAATCAAT	3821
QY	3994	GTACACTGAAGACCATGATTTGGATGTCCTCGAGAGAGACACATGAAGAATGATATGC	4053
Db	3822	TCACACCGACAGACAGATCTGGATGCTTGTGGAGAGACATTCCTGACCCAAAGCCATCTC	3881
QY	4054	TCACCTTACCTGTGAATCTGAATCTGAAGTGAAGTGGAGTAACTTCTGATGGG	4113
Db	3882	AAATTTCCCAAAAAGCTGCATCAATATGTGTGGAAAAAGGTGGAACTTCTGTGTGGG	3941
QY	4114	GGAAGGCGAGCTTGTGTGATTCATTCAGAGAGCCCTGCTCCGCCACTGTAAATTCGATTTT	4173
Db	3942	GGAAGGCGAGCTGTGTGATTCATTCAGAGAGAGCCCTGCTCCGCCACTGTAAATTCGATTTT	4001
QY	4174	AGATGAAGCCACAGCTCCATGTGACACAGACAGACATTAATTTGATTCAGAGACCATTCG	4233
Db	4002	CGATGAAGCCACAGCTCCATGTGACACAGACAGACATTAATTTGATTCAGAGACCATTCG	4061
QY	4234	AGAAGCATTTGACAGCTGTGCATTCGTCGACATTCGTCGACATTCGTCGACAGGCTCTAGG	4293
Db	4062	TGAAGCCCTTCCAGGGGTGACACCTGTGCTGATTCGTCGACATTCGTCGACAGGCTCTAGG	4121
QY	4294	CTCCGATAGGATTTGTGTGCTGTGCCCCAGGAGACAGTGTGTGATTTGACACCCCATTCGCT	4353
Db	4122	CTGTGACCGCATCTCTGTGTTATGTGGCATGTGGAAGTGTGTGATTTGATTCGCGGAGGT	4181
QY	4354	CTTCTGTCCACAGCATGTTCCGATTTCTATGTGCATGTTGCTGCTGC	4401
Db	4182	ACTGCGGAAGAACCTGGTGCATTTGTTGCGAGCCCTCATGAGCCACAGC	4229
RESULT	13		
LOCUS	AB012090		
DEFINITION	AB012090	765 bp mRNA	01-MAR-2000
ACCESSION	AB012090	Mus musculus mRNA for SMRP/MRP5, partial cds.	
VERSION	AB012090.1	GI:3551151	
KEYWORDS	SMRP/MRP5.		
SOURCE	Mus musculus		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Suzuki,T., Saseki,H., Kuh,H.J., Agui,M., Tatsumi,Y., Tanabe,S.,		

TITLE Terada, M., Saijo, N. and Nishio, K.
Detailed structural analysis on both human MRP5 and mouse mrp5 transcripts

JOURNAL Gene 242 (1-2), 167-173 (2000)

MEDLINE 20184734

REFERENCE 2 (bases 1 to 765)

AUTHORS Suzuki, T., Kuh, H. and Nishio, K.

JOURNAL Direct Submission
Submitted (09-MAR-1996) to the DDBJ/EMBL/GenBank databases.
Toshihiro Suzuki, National Cancer Center Research Institute,
Pharmacology Division: Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045,
Japan (E-mail: tssuzuki@cgan2.res.ncc.go.jp,
Tel: 81-3-3542-2511 (ex. 4451), Fax: 81-3-3542-1886)

Db	2142	CACATTACCTGGGGCCA--GGACGACGCCCTCCACACACTGAATGGCATCACCTTCCTCATC	219
Qy	1871	CAMAGGGCTAAACTGGTTGSGATCTGCGGCACTGTGGGAAGTGGAAAACTCTCTCATTT	1930
Db	2200	CCCGAAGAGTCTTTTGGTGGCGCGTGGGGCCAGAGTGGGGCTGGGAAATCTCTCCCTGGCTC	2259
Qy	1931	TCACGCCATTTTAGGCCAGATGACGCTTCTTAGAGGGGAGCATTTGCATACGTAGGAAACCTTC	1990
Db	2260	TCACGCCCTCTTGGCTGGATGGACAAAGTGGAGGGGGCACGTGGCTATCAAGGCGCTCGTG	2319
Qy	1991	GCTTATTTGGCCCCAGCAGGCGCTGGATCTCAATGCTACTGAGACACAACTCCTGTTTT	2050
Db	2320	GCCATATGCCCCACAGCAGGCGCTGGATTTCAAGATATGATCTCTCCGAGAAACATCTTTT	2379
Qy	2051	GGGAAGGAATATGATGAAAGAAAGATACAACTGTGTCTGAACAGCTGCTGTAGGCGCT	2110
Db	2380	GGATGTGACGTGGAGGAACCATATTTACAGTCCGATGATACAGGCTGTGCTCCCTCCCA	2439
Qy	2111	GACCTGGCCATTCTTCCACAGCCAGCTGCAGGGAGATTGGAGACGAGGCCAATCTG	2170
Db	2440	GACCTGGAATTCCTGCCCACTGGGATCGGACAGATGGGGAAGGAGGAGTGAACCTG	2499
Qy	2121	AGGGTGGGAGGCGCCAGAGATCACCTTGGCCGGGCGCTTATATAGTACAGAGCATC	2230
Db	2500	TCGTGGGGAACAGAAAGACGCGCTGAGGCTGGCCCGGCGCTACTCCAGCGCTACATT	2559
Qy	2231	TACATCTCTGAGCAGCCGCCCTCACTGCTTCATGATGCCCATGTGGGCAACACATCTTCAT	2290
Db	2560	TACCTTTCGATGATGCCCTCTCAGCAGTGGATGCCCATGTGGGAAACACATCTTTGAA	2619
Qy	2291	AGTCTTAT-----CGGAAACATCTCAAGTCCAGACAGTTCTTTGTATCCACACAG	2344
Db	2620	AATGTGATTTGGCCCCAAGGGGATGCTGAAGAACAAACACCGGATTTGGTACGCGACGC	2679
Qy	2345	TTACAGATCCGTTGGTGTGATGTAAGTATCTTATGAAGAGGCGTGTATTCGGAA	2404
Db	2680	ATGAGCTACTTGGCCGAGGTGGACGTCATCATCTGATATAGTGGGGCAAGATCTCTGAG	2739
Qy	2405	AGAGGCAACCCATGAGAACTGATGAAATTTAAATGGTGCATATGCTACCATTTTAAATAC	2464
Db	2740	ATGGGCTCTTACACAGGAGCTGTGGCTCGAGACAGGGCGCTTGCTGATGTCCTGGTACC	2799
Qy	2465	CTGTGCTGGGAGAGACACCCGCAAGTT---GAGATCAATTCAAAAAGAAACAGTGGT	2521
Db	2800	TATCCACGACAGCAGGAGCGAGGATGCGAGAGAGAAAGGGGGTCAACGGCGCTGACGCGT	2859
Qy	2522	TCACAGAAAGTGTCAACAAGACAGG-----GTGCTTAAACACAGATCAATTAAGAAAG	2573
Db	2860	CCAGGGAAGGAAGAAAGCAAAATGGAGAAATGGCATGCTGTACGCGACATGTGACGGAG	2919
Qy	2574	AAAAAGCAGTAAAGCCAGAGGAAGGGCACCTGTGTGCAGCTG-----	2615
Db	2920	CAATGTGAGAGACAGCTCACAGACTCTCTCTCTCTATAGTGGGGACATCAGCAGGACAC	2979
Qy	2616	AAGAGAAAGGGCAGGGTTCAGTCCCTGTGTCAGTATATGTGTCTACATCCAGGCTGCTG	2675
Db	2980	AACAGCACCACAGAACTGCAGAAAGCTGAGGCCACAGAAAGAAAGAACCTGGAAGCTGATG	3039
Qy	2676	GGGGG-----CCCTTGGATTCGAGGTTATTATGAGGCCCTTTATCTGTAATGTAGC	2728
Db	3040	GAGGCTGACAAAGGGGCGAGACAGGGCAGGTGAAGCTTCCGTACTCTGGGACTACATTGAG	3099
Qy	2729	AGCACCCCTTCACAGCACCTGTGTGTGAGTATCTGGATCAACCAAGAAAGGGGAAACCC	2788
Db	3100	GCCATGCGACTCTTTCACATCTCTCTCTCAGATCTTCTTTATGTGTAAACATGTGTC	3159
Qy	2789	ACTGTGACT-----CGAGGGAACGAGACCTGGGTGAGTGCAGCATGTAGAGACATCTT	2842
Db	3160	GCGCTGGCTTCCAATATTGGCTCACGCTCTGGACTGATGATGACCCCATGTCACAGGGGACT	3219
Qy	2843	CATTTGAGTACTATGTGCACACATCTTACGGCCCTCTCCATGGAGTCATCTGTATCTCTGAA	2902
Db	3220	CAGGAGACACGAAAGTCCGGCTGAGCGTGTATGAGACCCCTGGGATTTTCAAGAGGATC	3279

QY	2903	GCAATTCGAG-----AGTTGCTTTGTCACGAGGCGAGCTCTCGAGCTTCCTCCCGGGCTG	2956
Db	3280	GCGGTGTTTGACTACTTCATCGATGCGCGGTGCTCATCGGGGGGATCTTGAGCTTCCCGCTGTG	3339
QY	2957	CATGACGAGCTTTTCGAAAGATCCTTCGACGCCCTATGAACTTTTGTGACACGACCCC	3016
Db	3340	CACGTGACCTGCTGCACAGCATCTCGCGGTACACCATGAGCTCTTTGAGGAGGACCCC	3399
QY	3017	ACAGGGAGATTTCTCAACAGTTTTCACAAAGACATGAGATGAGTGTAGCTGCGCTGGCG	3076
Db	3400	AGTGGGAACCTGGTGAACCGCTTCTCCAAAGGAGCTGGACACAGTGAAGTCCATGATGCCG	3459
QY	3077	TTCCAGGCGCCAGATGTTTCACTCAGAACGTTATCCGTTGCTGTTGTTGGGAAATGAT	3136
Db	3460	GAGGTATCAAGATGTTTCATGAGGGCTCCGTTCACAGCTATTTGGGCGTCGATCGTTATC	3519
QY	3137	GCAGGAGTCTTCCCGTGGTTCTTGTGGCAGTGGGGCCCTTGTATCTCTTTTCACTG	3186
Db	3520	CTGTGGGCCACGCCCATCGCGCCGATCATCATCCGCCCTTGGGCTCATCTACTTCTTC	3579
QY	3197	CTGACATTTCTCCTCAGGGGTCGATTTGGGGAGTGAACGCTGTGACAAATATCACGAG	3256
Db	3380	GTCAGAGTTTCTACGTGGCTTCTCCCGAGCTGAAGCGCTCGAGTGGCTAGCGCC	3639
QY	3257	TCACCTTTCCCTCCACATCAGCTCAGCATACAGGGGCTTTGCCACCATCACGCTTAC	3316
Db	3640	TCGCCGGTCTATTTCCTCCATTTCAACGAGACCTGTCGGGGGTGAGGTCATTTGAGCTTTC	3699
QY	3317	AATTAAGGCGAGAGTTTCTGCACACATACAGAGACTCTGATGACACCAAGCTCTCT	3376
Db	3700	GAGGAGCAGGAGGCGTTCACTCCACAGAGTGAACCTGAAGTGGAGAGAACCGAGAGGCC	3759
QY	3377	TTTTTTTGTATTACGTGTGGATGCGGTGGGCGGTGCTGGCTGAGCTCATCAGATC	3438
Db	3760	TATTACCCAGCATCTGTGGCACAAGTGGCTGGCGCTGGAAGTGTGGGACAC	3819
QY	3437	GCCTCATCACACACAGGGGCTGATGTATGTTTATGCACGGGCAATTCGCCACCC	3498
Db	3820	TGCATCTTCTGTTTGTGCTGCCCTGTTTGGGCTGATCTCCAGGCACAGCTCAGTCTGGC	3879
QY	3497	TATCGGGGTCCGCACTCTTATGCTGTCACAGTTAAAGGGGCTGTTCAGTTTACGGTC	3556
Db	3880	TGTGTGGGCTCTCATGTGTCTTACTATTCAGAGTACACAGTACTTGAATGCGTGGTT	3939
QY	3557	AGACTGGCATCTGAGACAGAAAGCTCGATTCACTCGTGGGAGAGATCAATCACTACAT	3616
Db	3940	CGGATGTCATCTGAATGAAACCAATCATCTGTGGCGGTGAAGGCTCAAGAGATATCA	3999
QY	3617	AAGACTCTGTCTTGGAAAGCACTGCGACAAATTAAACAAGGCTCCCTCCCTACACGG	3676
Db	4000	GAGACTAGAA--GGAAGGCGCCCTGGCAAAATCCAGGAACAGCTCCGCCACGACGTGG	4056
QY	3677	CCCCAGGAGGAGAGAGTGGACCTTTGGAAGCGCAGAGATGAGTACCGAGAAACCTCCT	3736
Db	4057	CCCCAGTGGGCGAGTGGAAATTCGGAACACTAGCTCTCGCTACCGAAGAGCACTGGAC	4116
QY	3737	CTGCTCTTAAGAAAGTATCTTCCAGATCAAACTAAAGAGAAGTTGGCATTTTGGGG	3796
Db	4117	TTCTCTTCCAGGCAATCAATGTCCAGATCAAAAGGGGAGAAAGTGGGCACTCTGGGG	4176
QY	3797	CGGACAGATACGGAAGTCTCTGCTGGGAGTGGCCCTTCTCCGTCTGGTGGAGATTATCT	3855
Db	4177	CGGACGGGAGCTGGGAAGTGTCTCCCTGACCTTGCGTTTTCGGATCAAGAGTCTGGC	4236
QY	3857	GGAAGCTGATCAAGATTGATGAGAGTGAAGTCAATGATATTGGCTTGGCCGACTCCGA	3916
Db	4237	GAAAGAGATATCATCATCATGATGATCAATCAATCGCCAAATGGGCTCGACACACTCCGC	4296
QY	3917	AGCAAACTCTCTCATCTCTCTCAAGGCGCGGTCTGTAGTGGCAGCTGTGCATCAAT	3976
Db	4297	TTCAAGATCACATATCCCGCCAGAACCTGTTTTTTTTTGGGTTCCCTCGAATGAC	4356

Oy	3977	TTGAGCCCTTCAACACAGTACATCGAAGACAGATTGGGATGCCCTGGAGAGACAC	4036
Db	4357	CTGAGCCCATTCACGCCAGTACTCGGATGGAAGAGTGTGGACGTCCCTGGAGCTGCCAC	4416
Oy	4037	ATGAAAGATGTATTGCTCAGCTACCTCTGAAACTTGAATCTGAAGTATGATGACAAITGG	4096
Db	4417	CTGAAGACATCTCGTATCGTACGCCCTCTCCTGTGACAGCTGACCATCATATGTGACAAAGGCCGG	4476
Oy	4097	GATTAACCTTCTCAGTGGGGGAAAGGCGACCTTGTGATGATGACAGAGCCCTGCTCGGCCAC	4156
Db	4477	GAGAACTCTCAGTGGGGGAAAGGCGACCTTGTGATGATGACAGAGCCCTGCTCGGCCAG	4536
Oy	4157	TGTAAATTCCTGATTTTATGATGTAAGGCACACAGCTGCCTATGACACAGACAGACTTATG	4216
Db	4557	ACGAAAGATCCTTGTTGTTGATATGAGGCCACGCGACGCCGTGAGACTTGAAACGAGACACCTC	4596
Oy	4217	ATTCAAAGAGACCATCCGAGAGCAATTTGTCAGACTGTACATCCTGACCACTTGGCCCATCGC	4276
Db	4597	ATTCAGTTCACACATCCGAGACAGTTCCTGAGAGACTGCACCGTCTCTACCATCCGCCACCCGG	4656
Oy	4277	CTGACACAGGTTCTAAGGCTCCGATAGGATTTATGTGCTCTGGCCAGAGGACAGGTGGTAG	4336
Db	4657	CTCAACACCATCATGAGCTACACAAAGGGGTATGCTTGGACAAAGAGGAAATCCAGGAG	4716
Oy	4337	TTTGACACCCCATGGTCTCTTG	4360
Db	4717	TACGGCCGCCCATCGACCTCTCTG	4740

Search completed: November 9, 2001, 15:51:38
Job time: 12821 sec

Tue Nov 13 11:54:57 2001

us-09-528-031-1.rge

Page 33

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 13:04:42 ; Search time 278.09 seconds
(without alignments)
10944.080 Million cell updates/sec

Title: US-09-528-031-1
Perfect score: 4847
Sequence: 1 GGCTACTGCTCGGGACGTCG.....AAAAAAAAAGGGCGCCGC 4847

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq.0601.*
1: /SIDS1/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4847	100.0	4847	21	AAAA0481
2	4845.4	100.0	4847	19	AAV5682
3	4792.6	98.9	5838	20	AAZ30079
4	4792.6	98.9	5838	21	AAZ94745
5	4775.4	98.5	4781	22	AAC85287
6	1674.4	34.5	1761	18	AAZ94054
7	1054.2	21.7	2167	18	AAZ94055
8	511.2	10.5	5011	19	AAV31498
9	511.2	10.5	5011	20	AAV19818
10	511.2	10.5	5011	20	AAZ1977
11	511.2	10.5	5011	21	AAZ94741

12	511.2	10.5	5011	21	AAZ90193
13	511.2	10.5	5011	21	AAZ39556
14	506.4	10.4	4864	17	AAZ14911
15	506.4	10.4	4885	17	AAZ14910
16	506.4	10.4	5011	15	AAQ65377
17	506.4	10.4	5011	17	AAZ17173
18	506.4	10.4	5011	19	AAV31497
19	506.4	10.4	5011	20	AAZ19817
20	506.4	10.4	5011	20	AAZ1976
21	506.4	10.4	5011	21	AAZ60522
22	506.4	10.4	5011	21	AAZ90192
23	506.4	10.4	5011	21	AAZ39555
24	437.4	9.0	4231	20	AAZ30078
25	420.4	8.7	5889	19	AAV31499
26	420.4	8.7	5889	20	AAZ19819
27	420.4	8.7	5889	20	AAZ21978
28	420.4	8.7	5889	21	AAZ90194
29	420.4	8.7	5889	21	AAZ39557
30	404.4	8.3	437	21	AAZ02340
31	402	8.3	4834	17	AAZ14912
32	390.4	8.1	4918	18	AAZ94024
33	389	8.0	4669	17	AAZ14913
34	382.2	7.9	5586	18	AAZ94023
35	380.6	7.9	4864	21	AAZ94737
36	380.6	7.9	4868	19	AAZ26068
37	365.2	7.5	4931	19	AAZ29584
38	362.4	7.5	4509	20	AAZ30081
39	356.8	7.4	4762	18	AAZ94052
40	356	7.3	5079	20	AAZ30080
41	346	7.1	4877	16	AAZ01569
42	341	7.0	5110	16	AAZ01568
43	340.8	7.0	463	19	AAV5683
44	340.8	7.0	463	21	AAAA0482
45	340.8	7.0	463	22	AAC85288

ALIGNMENTS

RESULT 1	
AAAA0481	
ID	AAAA0481 standard; cDNA; 4847 BP.
XX	
AC	AAAA0481;
XX	
DT	23-NOV-2000 (first entry)
XX	
DE	Human MRP-beta cDNA.
XX	
KW	MRP-beta; multidrug-resistance associated protein; human; cytostatic;
KM	chemotherapy; cancer cell; drug discovery; cytotoxic; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	116..4429
FT	/*tag= a
FT	/Product= "MRP-beta"
PN	US6077936-A.
XX	
PD	20-JUN-2000.
XX	
PF	16-APR-1998; 98US-0061400.
XX	
PR	16-APR-1997; 97US-0843459.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
PI	Shyjan A;
XX	
DR	WPI; 2000-430613/37.
DR	P-PSDB; AAB10225.

Multidrug resistanc
Human MRP variant
cDNA encoding mult
cDNA encoding mult
Multidrug resistanc
cDNA encoding mult
Human multidrug re
Human multidrug re
Human multidrug re
Multidrug resistanc
Human multidrug re
Human multidrug re
cDNA encoding a hu
Murine multidrug r
Mouse multidrug re
Mouse multidrug re
Murine multidrug r
Human secreted pro
cDNA encoding mult
Rat canalicular mu
cDNA encoding mult
Human canalicular
Human ATP binding
Human secreted pro
Human canalicular
Homo sapiens sulph
cDNA encoding a hu
Human multidrug re
cDNA encoding a hu
Hamster sulphonylu
Rat sulphonylurea
Human multidrug re
Human MRP-beta cDN
MRP-beta probe, fo

Model multidrug-resistance associated polypeptide useful for improving the effectiveness of a chemotherapeutic regimen to eradicate multidrug-resistant transformed cells especially cancer cells -

Claim 6; Fig 1A-D; 43pp; English.

This invention describes a novel model multidrug-resistance associated polypeptide, MRP-beta (1) which has cytostatic activity. Inhibitors of (1) are useful for improving the effectiveness of a chemotherapeutic regimen to eradicate multidrug-resistant transformed cells, especially cancer cells, from the body of a mammal, preferably human. (1) is also useful for drug discovery, especially to the design of novel chemotherapeutic drugs that are cytotoxic to cells expressing (1). This sequence encodes the human MRP-beta protein which is described in the method of the invention.

Sequence 4847 BP; 1226 A; 1178 C; 1263 G; 1180 T; 0 other;

Query Match 100.0%; Score 4847; DB 21; Length 4847;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4847; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGCCTCATGCTCCGGAGGGTGTGAGGGGCTGGCGGCTGTCTCTGAGAGAGGGGGCGAG 60
1 ggcctcatgctccggaggggtgtgaggggctggcggctgtctctgagagagggggcgag 60
61 GAATTCGATGTAATACTAACAAGTCTGAGCCCTGGAGCCTCCACACAGAGAGATGAA 120
61 gaattcgaatgtaataactaacaagtctgagccctggagcctccacacagagagatgaa 120
62 gaattcgaatgtaataactaacaagtctgagccctggagcctccacacagagagatgaa 120
121 GGATATGACATAGAGAAAGATATATATATATATATATATATATATATATATATATAT 180
121 ggaatgacatagagaaagat 180
122 ggaatgacatagagaaagat 180
181 GAGAACACAGACATCTTGGAGCGCACAGACCCGTAAGATTCCAAATTCAGAGAACTCG 240
181 gagaacacagacatcttggagcgcacagacccgtaagattccaaattcagagaaactcg 240
241 ACCGTTGGAATGCCAATGATGCTTGAAGACAGACCCGAGAGCCCTCTCTTGA 300
241 accgttggaatgccaatgatgcttgaagacagacccgagagccctctcttga 300
242 accgttggaatgccaatgatgcttgaagacagacccgagagccctctcttga 300
243 accgttggaatgccaatgatgcttgaagacagacccgagagccctctcttga 300
301 TGCCCTCATGATCTCAGTCTGAGATCTGAGATGAGAGCATCCCAAGGAGAAATACCA 360
301 tgccctcatgatctcagctctgagatctgagatgagagcatcccaaggagaaatacca 360
361 TCATGCGTTGAGTCTTGAAGCCCATCCGACTACTTGGCAAACACAGCCCAAGTGA 420
361 tcatgcgttgagtcttgaagcccatccgactacttggcaaacacagcccaagtga 420
421 CAATGCTGGGGTCTTTCCTGATGATCTTTCGTTGCTTCTGCTGAGCGCTGAGC 480
421 caatgctggggtcttctctgatatgcttcttcgttgcttctctctctctctctctctct 480
481 CCACAGAGAGGGGAGCTCTCAATGGAAGAGTGTGTCTCTGTCACAAGAGAGTCTTC 540
481 ccacagagaggggagctctcaatggaagagtgtgtctctgtccaagagagagctcttc 540
541 TGACGTGAATGCGAAGAGCTAGAGAGCTGTGGCAAGAGAGCTGAATGAAGTTGGGCC 600
541 tgaagtgaatgcgaaagagctagagagctgtggcaagagagctgaatgaagttgggcc 600
601 AGAGCGTGGCTCCCTGGAGAGGGTGTGTGATGATCTTTCGCGCAGCAGAGCTCATCTTC 660
601 agagcgtggctccctggagaggggtgtgtgatgatcttctgcgcagcagagctcatcttc 660
661 CATGCTGCTGCTGATGATCAAGCAGCTGCTGCTTCAAGTGAAGCAGCTTCATGTTGAA 720
661 catgctgctgctgatgatcaagcagctgctgcttcaagtgaagcagcttcattggtgaa 720
721 ACACCTCTTGAGATATACCCAGGCAACAGAGCTTAACCTGCAATACAGCTTGTGTTAGT 780

721 acacctcttgagatataccagcaacagagcttaaccttcagtaagcgtgtgtgtagt 780
781 GCTGGGCTCTCTCTGACGGAATCGTGGGCTGTGCTGCTTGCATGACTGTGGGATT 840
781 gctgggctctctctgacggaatcgtgggctgtgctgcttgcattgactgtgggatt 840
841 GAATTCACGAACCGGTTCCGCTTGGCGGGGCGCATCTTAACATGGCATTTTAAAGAT 900
841 gaattcacgaacccggttccgcttggcggggcgcatcttaacatggcattttaaagat 900
901 CCTTAAGTAAAGAACTTAAAGAGAAATCCCTGGTGAGCTCATCAACTTGTCTCAA 960
901 ccttaagttaaagaaacttaaagaaatccctggtgagctcatcaacttgtctcaa 960
961 CGATGGCAGAGAAATGTTTGAAGCAGACGCCGTGGACGCTGCTGTGAGAGACCCGT 1020
961 cgatggcagagaaatgtttgaagcagacgccgtggacgctgctgtgagagaccgt 1020
1021 TGTTGCCATCTTAGCATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
1021 tgttgccatcttagcatgatgatatattattattattattattattattattattattatt 1080
1081 ATCAGCTGTTTATCTCTTATCCAGCAATGATGTTGATCATCAGCGCTCACAGCATA 1140
1081 atcagctgtttatctcttattccagcaatgatgttgatcatcagcgctcacagcata 1140
1141 TTTCAGAGAAATGCGTGGCGCCGACGATGAACGTGTCCAGAAATGATGAAGTTCT 1200
1141 tttcagagaaatgctggcgccgacgatgaacgtgtccagaaatgatgaagttct 1200
1201 TACTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1260
1201 tacttacaatataatataatataatataatataatataatataatataatataatata 1260
1261 GAAATCCGCGAGAGAGAGCGCTCGATATTGGAAGAAAGCCGGTACTTCCAGAGCATAC 1320
1261 gaaatccgcgagagagagcgctcgatattggaagaaagccggtaacttccagagcatac 1320
1321 TGTGGGTGTGGCTCCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
1321 tgtgggtgtggctccatttgt 1380
1381 GACCTGGGCTTCATGATGACAGCAGACAGCTTTCACACTGCTGACAGCTTCATTC 1440
1381 gacctgggcttcattgatgacagcagacagctttcacactgctgacagcttcatttc 1440
1441 CATGACTTGTGCTTGAAGTAACACCGTTTTCAGTAAAGTCCCTTCAGAGAGCTGAT 1500
1441 catgacttgtgcttgaagtaaacacggttttcagttaaagtccttcagagagctgat 1500
1501 GCGTGTGACAGATTTAAGAGTTGTCTTAATGAAAGAGTTTCACATGATTAAGAAACAA 1560
1501 gcggtgtgacagatTTAAGAGTTGTCTTAATGAAAGAGTTTCACATGATTAAGAAACAA 1560
1561 ACAGAGCAGTCTTCACATCAAGATAGAGATGAAGAAAGCCACTTGGATGGAGCTCTTC 1620
1561 acagagcagtcttcacatcaagatagagatgaagaaagccacttggatggagctcttc 1620
1621 ACCAGCAGTCTTCACATCAAGATAGAGATGAAGAAAGCCACTTGGATGGAGCTCTTC 1680
1621 accagcagtcttcacatcaagatagagatgaagaaagccacttggatggagctcttc 1680
1681 TTTCAGAGGCAAGAAAGAGAGTGAAGCAGCTGACGCACTGACATCAAGCGCTGCT 1740
1681 tttcagagggcaagaaagagagtgaaagcagctgacgcactgacatcaagcgctgct 1740
1741 GCGAGAGCAAGAAAGCCACTCTCTCTGAGAGAGAGAGCGGCGGCGGCAAGAGGA 1800
1741 gcgagagcaagaaagccactctctctgagagagagagcgggcgggcgagagga 1800
1801 AGAAGGCAAGCAATCCACTCGGGGCACTGCGCTTACAGAGGACATGCAACATGCA 1860
1801 agaaggcaagcaatccactcggggcactgcgcttacagaggacatgcaacatgca 1860

Db 1801 agaagcaagacacatccaccctgagccacctgcgtctacagagagacatgcacagcatcga 1860
 QY 1861 tcttgagatnccagaggggtaaacctggttggaattcggccgacatgtaggagagagaaaaac 1920
 Db 1861 tcttgagatnccagaggggtaaacctggttggaattcggccgacatgtaggagagagaaaaac 1920
 QY 1921 cttcttcattccagacattcttag 1980
 Db 1921 cttcttcattccagacattcttag 1980
 QY 1981 tggagaccttcgtctttagtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2040
 Db 1981 tggagaccttcgtctttagtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2040
 QY 2041 catctctgttggag 2100
 Db 2041 catctctgttggag 2100
 QY 2101 ccttgagagcctgacacgac 2160
 Db 2101 ccttgagagcctgac 2160
 QY 2161 agccaacctgagcagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2220
 Db 2161 agccaacctgagcagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2220
 QY 2221 cagagagccttctacatccctgagagacccctcagctgctttagatgccaatgagcagacaa 2280
 Db 2221 cagagagccttctacatccctgagagacccctcagctgctttagatgccaatgagcagacaa 2280
 QY 2281 catctctgttggag 2340
 Db 2281 catctctgttggag 2340
 QY 2341 ccaagttaacagatcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2400
 Db 2341 ccaagttaacagatcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2400
 QY 2401 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2460
 Db 2401 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2460
 QY 2461 ttaacctgttgcggag 2520
 Db 2461 ttaacctgttgcggag 2520
 QY 2521 ttcacagaagaagatccacaagacaaaggttccctaaacaagatacaataaagaagaaaaagc 2580
 Db 2521 ttcacagaagaagatccacaagacaaaggttccctaaacaagatacaataaagaagaaaaagc 2580
 QY 2581 agtaaacgcaag 2640
 Db 2581 agtaaacgcaag 2640
 QY 2641 cttgagcagatatactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2700
 Db 2641 cttgagcagatatactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2700
 QY 2701 tatgagccttcttcaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2760
 Db 2701 tatgagccttcttcaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2760
 QY 2761 cttgagcagatatactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2820
 Db 2761 cttgagcagatatactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2820
 QY 2821 ttcacagaagaagatccacaagacaaaggttccctaaacaagatacaataaagaagaaaaagc 2880
 Db 2821 ttcacagaagaagatccacaagacaaaggttccctaaacaagatacaataaagaagaaaaagc 2880
 QY 2881 ggcagcagatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2940
 Db 2881 ggcagcagatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2940

QY 2941 agcttccctccggcctgac 3000
 Db 2941 agcttccctccggcctgac 3000
 QY 3001 ttttgag 3060
 Db 3001 ttttgag 3060
 QY 3061 tggagaccttcgtctttagtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3120
 Db 3061 tggagaccttcgtctttagtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3120
 QY 3121 cttgagcagatatactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3180
 Db 3121 cttgagcagatatactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3180
 QY 3181 catctctgttggag 3240
 Db 3181 catctctgttggag 3240
 QY 3241 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3300
 Db 3241 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3300
 QY 3301 ccaaccttccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3360
 Db 3301 ccaaccttccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3360
 QY 3361 tgaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3420
 Db 3361 tgaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3420
 QY 3421 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3480
 Db 3421 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3480
 QY 3481 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3540
 Db 3481 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3540
 QY 3541 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3600
 Db 3541 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3600
 QY 3601 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3660
 Db 3601 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3660
 QY 3661 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3720
 Db 3661 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3720
 QY 3721 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3780
 Db 3721 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3780
 QY 3781 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3840
 Db 3781 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3840
 QY 3841 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3900
 Db 3841 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3900
 QY 3901 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3960
 Db 3901 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3960
 QY 3961 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 4020
 Db 3961 ggaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 4020

```

OY 4021 CCTGGAGAGGACACATGAAAGATGATTGCTCAGCTACCTCTGAACCTTGAATCTGA 4080
    |||||||
DB 4021 cctggagagagacacacatgaaagtatgtatgtctcagctacctctgaaactgaaactgaa 4080
OY 4081 AGTATGAGAGATGGGATTAATCTCTCAGTGGGGGAACGGACGCTGTTGCATAGCTAG 4140
    |||||||
DB 4081 agtataaggagaaatgggaaatcctcagtcgggaaagagcccttgatgaaactgaaactgaa 4140
OY 4141 AGCCCTGCTCCGACAGTGAAGATTCTGATTTTGTAGTAAAGCCACAGCTGCATGGACAC 4200
    |||||||
DB 4141 agccctgtcccgacagtgtaagattctgatttagatgaagcaagctgcacagtcgaaacac 4200
OY 4201 AGAAGACAGACTTATGATTGATTCAGAGACCATCCGAGAAACATTTGACAGCTTACCATCT 4260
    |||||||
DB 4201 agagacagacttattgtattcaagagacatccgaaagcaacttgcagactgtacacatgct 4260
OY 4261 GACCATTCGCCATGCGCTGCACAGGTTCTAGGCTCCGATAGGATTATGAGTGGGCCCA 4320
    |||||||
DB 4261 gaccattgcccacatcgccgacacaggtctcaggtccgataagattatgctggtgcccga 4320
OY 4321 GGGACAGGTGGTGAAGTTTGAACCCCATCCGTCCTCTGTCACAGACAGATTCGCCGAT 4380
    |||||||
DB 4321 gggacaggtgtgtgagatttgacaccccacatcggtctctctgtccaaagcagttcccgatt 4380
OY 4381 CTATGCCATGTTTGTGCTGTCAGAGAACAGCTGCTGTCAAGGGCTACTCTCTCTGT 4440
    |||||||
DB 4381 ctatgacatgtttgtcgtcgcagagaaaggtcgtctcaagggcttgaacctctctcgtt 4440
OY 4441 TGACGAGTGTCTTCTTCTTGTAGACATTCGATTCCTGCTGGGGGGGGCCCTTCATC 4500
    |||||||
DB 4441 tgaagaagctctcttctcttcttgagatgcatctccctgctgggggggcccctctatc 4500
OY 4501 GCGTCTCTCTACCGAAACCTTGCTTGTGATTTATCTTTCGACAGCAGTTCGGAT 4560
    |||||||
DB 4501 gcgctctctctacgaaacctgtctctcgaattatctcttcgacagcagttccgatt 4560
OY 4561 TGGCTGTGCTTTCACCTTATAGGAGAGCATATTTGATTTATTTATTTATTCATAT 4620
    |||||||
DB 4561 tggctgtgttctcactctttagagagatcatattgtatattgtattatctcatat 4620
OY 4621 TCATGTAAACAAATTAATTTTGTCTTCTTATTAATGACACTTAAAGGTTTCAGGAAACCGT 4680
    |||||||
DB 4621 tcatgtaaacaaatttagttgttcttcttaattgcaactctaaaggttcagggaacctgt 4680
OY 4681 TATTATTAATTTATTCAGAGGCTTATGAACTTATACCTGTAGCTATATCTATATAT 4740
    |||||||
DB 4681 tatataattgtatcagaagcctataaagacttatacgtgtagctataatctatata 4740
OY 4741 AATTCTGATACATAGCTATTTTACAGTAAATGTAAAGCTTATTTTATTTAAAT 4800
    |||||||
DB 4741 aattctgataacatagcctataattacagtgaaatgtaagctgttatttattatataaata 4800
OY 4801 AAGCACTGCTCTAAATAAAAAAAAAAAAAAAAAAAGGGCGCGC 4847
    |||||||
DB 4801 aagcactgtgtctaaataaaaaaaaaaaaaaaaaaaggcgccgc 4847

```

RESULT 2

AAV5682 standard; cDNA; 4847 BP.

AAV5682;

02-FEB-1999 (first entry)

Human multidrug resistance-associated protein cDNA.

Multidrug resistance-associated protein; MRP-beta; human; cancer;

tumour; chemotherapy; therapy; ss.

Homo sapiens.

Key Location/Qualifiers

```

FT CDS 116..4429
FT FT /*tag= a
FT PN MO9846736-A1.
FT XX 22-OCT-1998.
FT PD 16-APR-1998; 98WO-US07673.
FT PF 16-APR-1998; 97US-0843459.
FT PR 16-APR-1997; 97US-0843459.
FT PA (MILL-) MILLENNIUM PHARM INC.
FT XX Shyjan A:
FT PI WPI: 1998-568724/48.
FT DR P-PSDB; AAM80597.
FT PT New isolated multidrug resistance-associated polypeptide - used to
PS develop products for modulating multidrug resistance, particularly
PS for reducing resistance of tumours to chemotherapeutic drugs
PS Claim 1; Page 54-59; 93pp; English.
CC This nucleotide sequence codes for novel human multidrug resistance
CC associated protein (MRP-beta, see AAM80597), over-expression of which
CC is thought to be associated with the emergence and/or persistence
CC of a multidrug-resistance phenotype in transformed mammalian cells,
CC including carcinoma and adenocarcinoma cells. A unique fragment
CC (see AAV5683) of the MRP-beta gene was identified by computer
CC assisted searching of a nucleic acid database corresponding to a
CC human endothelial cell (HMEC) expression library. It was used
CC to screen the HMEC expression library. This yielded a 4.78 kb
CC clone, designated fohd013a05m (deposited as ATCC 98409). Two
CC independent cDNA clones comprising approximately 60 residues
CC upstream from fohd013a05m were isolated by hybridisation screening
CC of human brain and liver cDNA libraries using a probe from the 5'
CC end of fohd01305m. The presented sequence of MRP-beta cDNA
CC comprises fohd01305m and 66 additional upstream nucleotides. The
CC invention provides compositions and methods for improving the
CC effectiveness of chemotherapeutic regimens to eradicate
CC multidrug-resistant transformed cells from the body of a mammal,
CC especially a human. The disclosed compositions include MRP-beta
CC nucleic acids, including probes and antisense oligonucleotides
CC (see also AAV5684-88), MRP-beta polypeptides and antibodies,
CC MRP-beta expressing host cells, and non-human mammals that are
CC transgenic or nullizygous for MRP-beta. The disclosed methods
CC include methods for attenuating aberrant MRP-beta gene expression,
CC protein production and/or protein function, and for improving the
CC effectiveness of chemotherapy for a mammal afflicted with a
CC multidrug-resistant tumour, wherein the tumour is of mammary,
CC respiratory tract, urogenital tract, endocrine system or immune
CC system origin. In addition, methods are disclosed for identifying
CC and using a modulator, such as an inhibitor, that is cytotoxic to
CC cells expressing MRP-beta.
XX Sequence 4847 BP; 1225 A; 1179 C; 1263 G; 1180 T; 0 other;
SQ

```

Query Match 100.0%; Score 4845.4; DB 19; Length 4847;

Best Local Similarity 100.0%; Pred. NO. 0;

Matches 4846; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 GGCATATGCTCGGAGAGCTGTGGTTGAGCGGTGGCGGCTTCTCGAGACAGGGGCGCAG 60
    |||||||
DB 1 ggcataatgctcgaggagcgtgtgtgagcgcgtgcggttctctgagcagcgcgag 60
OY 61 GAATTCGATGTGAAGTAACAGTCTGTGAGCCCTGGAACCTCCACTCAGAGAAATGAA 120
    |||||||
DB 61 gaattctgatgtgaagtaaacagctctgtgagccctggacctccactcagaagaatgaa 120
OY 121 GGATATGACATAGGAAAGAGTATATCATCCCAAGTCTCTGGGTATAGAAATGTGAGGGA 180
    |||||||

```


D 121 gatatcgacatagyaagaagatatalcatccccagctctggtatagaaagtgtagggga 180
Q 181 GAGAACCGACGCTTGTGGAGACACAGACCGTGAAGATTCCAGTTCCAGAGAACTCG 240
D 181 gagaacccgacctctcgtgagcagacagagacggtgaagattccaaagttcagagaactcg 240
Q 241 ACCGTTGGAAATGCCAAGATGCTTGGAAACAGACAGCCGAGCCGAGGGCCCTTCTCTTGA 300
D 241 accgtttgaaatgcccagaatgctcttggaaacagccgcccagcggagggctctctcttga 300
Q 301 TGCCCTCATGCAATTTCTCAGTCTCAGAAATCCGAGTAGAGAGACATCCCAAGGAAATACCA 360
D 301 tgcctccatgatcttcagctcagatcccttgatgagagatccccaagggaagtacaa 360
Q 361 TCATGCGTTGATGCTGTGAGAGCCATCCGAGACTTGTGCAAAACACAGACCCAGTGGGA 420
D 361 tcaatgcttgcgtcctcgaagcccaaccgagctacttgcgaacacccgacccagtgga 420
Q 421 CAATGCTGGGCTTTTTCCTGATGACTTTTTCGTGCTTCTTCTGCCCCGTGTGAC 480
D 421 caatgctgggcttttctctctgtatgacttttcgtgcttctctctctgcccgtgtgac 480
Q 481 CCACAGAGAGGGGAGCTCTCATGTGAAGAGCGTGTGCTCTGTCCAAACAGACGACTTTC 540
D 481 ccaacaagaaggggagctcctcaatggaagacgtgtgctctgtccaaagcacagactcttc 540
Q 541 TGAGCTGAAGTACGACAGACTAGAGAGACTGTGCAAGAGAGCGTAGTGAAGTGGGCG 600
D 541 tgaagctgagcagcaagaagactagagaagactgtgcaagaagaagcgaagatgtgggc 600
Q 601 AGAGCGTCTTCCCTGCGAAGGGGTGTGTGATCTTCTGCCCCAGACAGGCTCATCTGTC 660
D 601 agaagctgtctccctcggaagggtgtgtgtgactcttcgcgcacagactcatctctgc 660
Q 661 CATGCTGTGCTGATGATCAGCAGCAGCTGCTGCTTCACTGAGTGAACAGCTTCATGCTGA 720
D 661 catgctgtgctgtatgatacagcagcagctgtgctctcagtggaaccaagcctcatagtgaa 720
Q 721 ACACCTTGGAGTATACCCAGGCAAGAGACTTAACCTGCAAGTACAGTGTGTTGAT 780
D 721 acaactcttgagatatacccaagcaacagagctaaacctgacagatacagctgtcttaag 780
Q 781 GCTGGGCTCTCTCTGAGGGAATGCTGCGGCTTGTGCTGCTTGCACACTGACTTGGGCAAT 840
D 781 gctgggctctctccctcgaggaatcggtgtgtctgtgcctgacactgagacttgggcaat 840
Q 841 GAATTACCGAACCAGGTGCTGCTGCGGGGCGCATTCCTAACATGGCATTTAAAGAAAT 900
D 841 gaattaccgaaccaggtgtcgtcttgcggggccatcccaacatgycattlaagaagat 900
Q 901 CCTTAAGTTAAAGACATTTAAAGAGAATCCCTGGGTGAGTCAATCATTTGCTCCAA 960
D 901 ccttaagttaaagacatlaaagaagaatccctgggtgagctcaacaatctgtcccaa 960
Q 961 CGATGGGAGAGAAATGTTTGAAGGAGCAGCCGTTGGACGCTGTGCTGAGAGAGCCGT 1020
D 961 cgaatggagagaagaatgtttgagcgacgacgctgtgcagccgctgctggagagccgt 1020
Q 1021 TGTTCGCATCTTAGGAGATTTATATATGTAATTAATCTGGAGCAACAGGCTTCTGTGG 1080
D 1021 tgttgcctcttaagcagatcttaatacttaattctcggagcaacaagcttctctg 1080
Q 1081 ATACACTGTTTATCTCTTTTACCCAGCAATGATGTTTGCATCAGGCTTCACAGCATA 1140
D 1081 atcacctgtcttctaccccttctacccagcaatgattgttgcatacaggtctacagcata 1140
Q 1141 TTTTCAGGAGAAATGCTGCGCCGACGAGATGAAGTGTGCAAGATGAAGTGTCT 1200
D 1141 ttctcagagaaaatgctggtgcgcgacgagatgaagctgtccagaagatgaatggaagttct 1200
Q 1201 TACTTACATTAATTTATCAAAATGTATGCTGTGCTCAAGCATTTTCTCAGAGTGTTC 1260
D 1201 tacttaacatlaattatcaaaaatgattgctgtggtcaaacatcttctcagagtggttca 1260

Q 1261 GAAATCCGCGAGAGAGAGCGCTCGGATATTGAAAAAGCCGGGTACTTCCAGAGCATCAC 1320
D 1261 gaaatccgagagagagcgctcggatattgaaaaagccgggtacttccagagcatcac 1320
Q 1321 TGTGGTGTGCTCCCATTTGTGTGTGATGTCACAGCGTGTGACCTTCTCTGTTCAAT 1380
D 1321 tgtgggtgtgctcccatgtgtgtgtgattgacagcggtgtgacctctctctgctcatat 1380
Q 1381 GACCTGGGCTTCGATCTGACAGCAGCAGGCTTTCACAGTGTGTGACGTTCAATTC 1440
D 1381 gacctgggcttcgactcgaagcagcagcagcttccacagtggtgacagcttccaatc 1440
Q 1441 CATGACTTTTGTGGAAGTAAACCGTTTCAATGAAAGTCCCTGCGAAGCCTCAGT 1500
D 1441 catgacttttgtggaagttaaacctgttcaatgaaagtccctgcaagccctcagt 1500
Q 1501 GCGTGTGACAGATTTAAGAGTTGTTCTAATGGAAGAGTTCAATGATTAAGAAACA 1560
D 1501 gcggtgtgacagatttaagagttgttctaattggaagagttcaatgattaagaaaca 1560
Q 1561 ACCAGCCAGTCTCATATCAATAGATGAGATGAAAAATGCCACTTGGCATGGGACTCTC 1620
D 1561 accagccagctctcatcaatagatagatgaaaaatgccacttggcatgggactcttc 1620
Q 1621 CCACCTCAGTATCCAGAACTCGCCAGCTGACCCCAAAATGAAAAAGACAAGAGGC 1680
D 1621 ccacctcagtatccagaatcgcgccaaagctgcaccccaaaatgaaaaagacaagaggc 1680
Q 1681 TTCAGAGGCGCAAGAAAGAGAGGTAGGCGAGCTGACGCACTGAGCATCAGCGGTCT 1740
D 1681 ttccaggggcaagaagaaggtgagcagctgcagcgactgagcatcagcggtgtct 1740
Q 1741 GGCAGACGAAAGGCGCACCTCTCTCTGACAGTACAGAGCGGCCCACTGCCAAGAGA 1800
D 1741 ggcagacgaaaggcgccacctctctctgacagtgacgagcgcccaagctccgaagaaga 1800
Q 1801 AGAAGCAAGACATATCCACTGCGCCACCTGCGCTTACAGAGAGACACTCAGAGATCA 1860
D 1801 agaagcaagacatattccactgcccacactgcgcttacagagaaactcagagcatcga 1860
Q 1861 TCTGAGATCCAAAGAGGTAACTGTTGAATCTGCGGCACTGTGGAGTGGAAAAAC 1920
D 1861 tctgagatccaaagaggttaactgtttgaaatctgcgcaggtgtggaggtgaaaaac 1920
Q 1921 CTCTCATATTTAGCCATTTTGAAGCAGATGACGCTTCTAGAGGCGAGCATTCGATCG 1980
D 1921 ctctcatatttagccattttaggcagaatgacgtcttagagggcagcatgtgcaatcag 1980
Q 1981 TGGAACTTCGCTTATGNGGCCAGAGGCTGGAATCCCTCAATGCTACTGAGAGACA 2040
D 1981 tggaaacttcgcttattgnggccagagcctggaaatccctcaatgctactgagagaca 2040
Q 2041 CATCTGTTTGGGAAGGAATATGATGAAGAAAGATACAATCTGTGCTGACAGCTGCTG 2100
D 2041 catctgtttgggaaggaaatgatgaagaaagatacaactctgtgtgaacagctgctg 2100
Q 2101 CCTGAGGCTGACCTGCGCATTTCTTCCAGACGACCTGACGAGAGATTGAGAGCGAGG 2160
D 2101 cctgaggctgacactgcccatttccccaagctgacagctgagagattgagagcgaagg 2160
Q 2161 AGCCAACTGACGAGGGGGGCGACGCGCAGAGATCAGGCTTGGCCGGGCTTGTATAGTA 2220
D 2161 agccaaactgacgagggggcgacgcgccagagatcagccttgcggggctgtatagtgaa 2220
Q 2221 CAGGAGCATCTACATCTTGAAGAGACCCCTCAGTGTGCTTGAATGCCATGTGGCAACCA 2280
D 2221 caggagcatctacatcttgaagagacccctcagtgcttgaagtcccatgtgtggaacca 2280
Q 2281 CATCTTCAATAGTCTATCCGGAACATCTCAAGTCCAAAGACAGTTCTGTGTTTACCCA 2340
D 2281 catcttcaatagtgctatccggaacatctcaagtcgaagacagtgctgttcttaccca 2340

2241 CCAGTACAGTACCTGGTGTGACTGTGATGAAGTACTCTTCATGAAAGAGGGCTGTATTAC 2400
2241 ccaagtacagtaacctggtgtgactgtgataagtgatctcttcataagaagggctgtattac 2400
2401 GGAAGAAGGACCCATGATGAGAACTGATGAATTTAAATGGTGCATGCTATGCTACATTTTAA 2460
2401 ggaagaaggaacccatgatagaaactgatgaatttaaattgggtgcatagtctacattttaa 2460
2461 TAACTTTGCTGGGAGAGACACCGCAGTTGAGATCAATTCAAAAAGAAACAGATGG 2520
2461 taactttgctgggagagacacccgagttgagatacaattcaaaaaagaaacagatgg 2520
2521 TTCACAGAGAAGTGCACAGACAAAGGCTCTTAAACAGAGATCAATTAAGAAGAAAAAGC 2580
2521 ttcacagaagaagtgcacaagacagggtcccttaaaacagatcaataagaagaagaaaaagc 2580
2581 AGTAAACCCAGAGAAAGGACGCTTGTGACGCTGGAAGAGAAAGGCGAGGTTACAGTCC 2640
2581 agttaaaccagaaaggaagggcagctgtgcagctggaaagaaagggcaggttcaagtcc 2640
2641 CTGGTCAGTATATGTTGTCATCCACAGGCTGTGGGGCCCTTGGCATCTGTTAT 2700
2641 ctggtcagtatatgtgtgtctacataccagctgtcgtggggcccttggcattccgtttat 2700
2701 TATGGCCCTTTTATGCTGAATGTAGCAGACCCGCTTCAGACCTGTGGTGAAGTTA 2760
2701 tatggcccttttcatgtctgaatgttagagcaagcccttcagacactgtgtgtgaagtta 2760
2761 CTGGATACAGAAAGGCGGGAACCACTGTGACTGAGGGAACGAGACTGGGTGAG 2820
2761 ctggatcaagaagaagcgaggaaacacacgtgactcgaaggaaagagactcgggtgag 2820
2821 TGAGAGCATGAGAGACATCTCATATGACATATGACAGATCTACGCTCTCCAT 2880
2821 tgaagacatgaagagacaatctcatatgcaagtactatgccaagcatcgaagccctccat 2880
2881 GGCAGTATGCTGATCCTGAAAGCATTTCCAGAGTGTCTTTGTCAAAGGACGCTGCG 2940
2881 ggcagtatgtctgactcctgaagcaatcgaagagatgtcttctgaaggcacgtgcg 2940
2941 AGCTTCCCGGCTGCATGACGAGCTTTTCCGAAGATCCTTGAAAGCCCTATGAGTT 3000
2941 agcttcccggtctgcatgacgaagcttccgaagatctccgaagccctatgaagtt 3000
3001 TTTTGACAGACCCCAAGGAGGATTTCCACAGGTTTTCCAAAGACATGATGAGT 3060
3001 ttttgacagaccccaaggagatcttcaacagtttccaagaacatgagtgaagt 3060
3061 TGAGTCGCGCTGCGGTTCCAGCGCGAGATGTTATCCAGAACTTATCCTGGTGTCTT 3120
3061 tgaagtcgctgcgtccgtccagagatgtltaaccagaagctltaacctgtgtctct 3120
3121 CTGCTGGGATGATCCGACAGAGTCTCCCGTGGTTCCTTGTGACAGTGGGGCCCTTGT 3180
3121 ctgctgggaatgatccgacagagcttcccggtgttctctgtgagatggggcccttgt 3180
3181 CATTCTTTTCACTCTGCAATGTGTCACAGGGTCTTGATTTGGGAGCTGAAGCTCT 3240
3181 catctctttcagctctgcatctgtctcgaaggtctcgtlctcgagagctgaagcgct 3240
3241 GGACAAATACAGGAGTACACTTTCCCTCTCCACATACAGTCCAGCATACAGGGCCTTGC 3300
3241 ggaacaatatcagagatcaacttctctccacatacagctccagacatacagggccttgc 3300
3301 CACCATCAGCGCTACATATAAGGGAGAGTTCTGCACAGATACAGAGAGCTGTGA 3360
3301 caccatcacaagctacaataaaggagagttcttcgaacagatacacaagagctcgtgga 3360
3361 TGACAAACAAAGCTCTTTTGTGTTTACGTGTGCAATGCGGTGGCTGCTGTGGGCT 3420
3361 tgacaaacaaagctcttcttcttgaagcatgcaatccctcgtggcgctgctgtgggct 3420
3421 GGACCTATACGATCGCCCTCATACACACAGGGGCTGATGATGTTTATGACAGG 3480

3421 ggaacttaacagatcgcctcccaacacacacagggctgtgactgtcttcttgcacgg 3480
3481 GCAGATTCGCCAGCGTATGCGGGTCTGCCATCTTATCTGTCGACGTAATGCGGCT 3540
3481 gcaatlcgcccaagcctatgctggtctccgcacatctctatgtcttccagttaaagggct 3540
3541 GTTCCAGTTTACGGTACAGTGCAGTCTGAGACAGAAAGCTTGATTCACCTGGTGAAG 3600
3541 gtccagtttaacggttcaagactcgtgcatctgagacagaagctcgtatccactcgtggagag 3600
3601 GATCAATACATCAATTAAGATCTGTCTTGGAGGACACTCCGAGATTAAGAACAGGC 3660
3601 gatcaatacactacaattaagactcgtctcttggaaagcactgtccagaatlaagaacaagc 3660
3661 TCCCTCCCTGACTGCGCCCGAGAGGAGGTGACCTTTCAGAACGACGATGAGGTA 3720
3661 tccctccctgactgccccagagagggaggtgactcttgaagaacgagagatgaggtla 3720
3721 CCGAGAAAACCTCCCTCTCGTCTTAAGAAAGTATCCTTCACGATCAAACTTAAGAA 3780
3721 ccgagaaaacctccctctcgtcttaagaagttactcttcaacgatacaaaactaaagaa 3780
3781 GATTGGCATTTGGGGCGGACAGGATCAGGAAAGTCTCGCTGGGGATGGCCCTTCCG 3840
3781 gattggcatttggggcggaacagatcaggaagtctcgtggggatggcccttccg 3840
3841 TCTGTGAGATTATCTGGAGGCTGCATCAATGATGATGAGATGAGATCAATGATTTGG 3900
3841 tctgtgagattatctggagggctgcatacaagaatgagatgaggtgagatcagttatgg 3900
3901 CTTGCGCAGCTCCGAAGCAAACTCTATCATTTCTTCAGAGCGGCTGCTCACTG 3960
3901 cttgcgagctccgaagcaaaactctcatctatcttcaagaagcggtgcttcaagtg 3960
3961 CACTGTCAGTCAATTTGGACCCCTTCAACGATACGTAACGATTTGGGATTC 4020
3961 cactgtcagatcaaatltggaccccttcaacagtaacgtgaagacagatctggagatgc 4020
4021 CTTGGAGGACACACATGAAAGATGATTTGCTACAGTCACTGAAACTTGAATCGA 4080
4021 cttggagggacacacatgaaagatgtatgtctcagctactcttgaacttgaatctga 4080
4081 AGTATGAGAAATGGGATTAATCTCTCAGTGGGGGAACGGAGCTTGTGCTATGACTAG 4140
4081 agtatggaagaatgggataactctcagttggggaaagagctctgtgcatagtag 4140
4141 AGCCCTGCTCGCCACGTGAATTTGATTTTGAAGAACCCACAGCTGCATGAGCAC 4200
4141 agccctgctccgacacgtgaatttgaatttagatgaagcacagctgcacatggagac 4200
4201 AGAGACAGACTTATTTGATTAAGAGACCATCCGAGAACATTTGACAGTGTACATGCT 4260
4201 agagaagacttatgtatgaagagacacatccgagaagcatltgcagacgttaccatgct 4260
4261 GACCATTCGCATCGCTGCACACAGTTCAGCTCCGATAGGATTTATGCTGGGCCCA 4320
4261 gaccattcgcatacgctgcacacagttctaggtctcgataggtltaatgtgctggccca 4320
4321 GGGACAGGTGAGGATTTGACCCCATCGGCTCTGTCACAGCAGATTTCCCATTT 4380
4321 gggacaggtgaggaatttgaccccatcggtctctgcccacagacatctccagatt 4380
4381 CTATGCCATGTTTCTGCTGCAGAGAACAAAGTGCTGTCAAGGGCTACTCTCCCTGT 4440
4381 ctatgcatgttctgtctgtgagaaaggtcgtctgaaggctcgaatccctccctgt 4440
4441 TGACGAAGTCTCTTTTCTTATGACATTTGCCATTCCTGCTGGGGGCGCCCTTATC 4500
4441 tgacgaagtctcttcttctttagagcatgtccatccctcgtgggggcccccttacc 4500
4501 GCGTCCCTACGGAACCTTGCTTCTGATTTTATCTTGCACAGAGTTCGGAT 4560

Db	4501	gcgcctccctacgcgaaccttgctcttcctcgatttctcttcgcgaagcagtlccggtat	4560
Oy	4561	TGCGCTTGTCGTTCACATTTTATAGGAGATCATATTTTGTATTGATTATTCATAT	4620
Db	4561	tgctcttgctgcttcctactctttaggagagatcattatcttgatttcttccat	4620
Oy	4621	TCATGTTAAACAAATTTTAGTTTTTGTCTTATATTGCACCTCTAAAGTTACAGGAACCGT	4680
Db	4621	tcacgttaacaaaatttagtttcttgccttaattgcaactaaaggctcaaggaaaccgt	4680
Oy	4681	TATATATATTGTATCAGAGCGCTTATATGAAGCTTTATAGCTATATATCATATAT	4740
Db	4681	tattataattgtacagagagcctaataagccttctacgctgtagctatactataat	4740
Oy	4741	AATTCGTGATACCTCATATTTTACAGTGAAAATGTAAAGCTGTTATTATTATAAT	4800
Db	4741	aattcgtlacatagcctatactattacagtgaaatgtaagctgcttattctataataat	4800
Oy	4801	AAGCAGCTGCTATAAAAAAAAAAAAAAAAAAGCGCGCGCCG	4847
Db	4801	aagcactgtgctaaaaaagggcgcgcg	4847
RESULT 3			
ID	AA230079	standard; cDNA; 5838 BP.	
XX	AA230079;		
XX	26-JAN-2000	(first entry)	
DE	cDNA encoding a human MPR-related ABC transporter designated MOAT-C.		
XX	Human; MPR-related ABC transporter; MOAT protein; MOAT-C;		
KW	MOAT mediated transport; anticancer drug sensitivity;		
KW	transporter mediated cellular efflux; anticancer; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	126..4439	
FT		/*tag= a	
FT		/product= "MOAT-C"	
FT		/note= "MPR-related ABC transporter"	
XX	W09949735-A1.		
XX	07-OCT-1999.		
XX			
PF	26-MAR-1999;	99MO-US06644.	
XX			
PR	27-MAR-1998;	98US-0079759.	
PR	03-AUG-1998;	98US-0095153.	
XX	(FOXO-) FOX CHASE CANCER CENT.		
XX			
PI	Kruh G, Lee K, Belinsky M, Bain L;		
XX			
DR	WPI: 1999-610812/52.		
DR	P-PSDB: AAY43542.		
XX			
PT	New transporter gene useful for screening for anti-cancer drugs		
XX			
PS	Claim 11; Page 134-135; 153pp; English.		
XX			
CC	The present sequence encodes a human MPR-related ABC transporter (MOAT)		
CC	protein, designated MOAT-C. The protein comprises a multi-domain		
CC	structure including a tandem repeat of nucleotide binding folds		
CC	appended C-terminal to a hydrophobic domain, having Walker A and B ATP		
CC	binding sites and several potential membrane spanning domains. The MOAT		
CC	nucleic acids are useful for screening a test compound for inhibition of		
CC	MOAT mediated transport, indicated by restoration of anticancer drug		
CC	sensitivity, which in turn causes a reduction of transporter mediated		

CC cellular efflux of anticancer agents. MOAT DNA or RNA may be used as
CC probes to detect the presence or expression of genes encoding MOAT
CC proteins. Anti-MOAT antibodies are useful for detecting and quantitating
CC MOAT proteins.

sequence 5838 BP; 1422 A; 1432 C; 1509 G; 1475 T; 0 other;

Query Match	98.9%	Score 4792.6	DB 20	Length 5838
Best Local Similarity	99.8%	Pred. No. 0		
Matches 4809	Conservative	0	Mismatches 9	Indels 1
				Gaps 1

QY	1	GGCTATCTCTGGGAGAGCGTGGTTTACGCGCGTGGCGGCTGTCTCTGGAGCAGGGGGCGACG	60
DB	11	ggctaatcgtctgggagcgctgtgtgagcgctcgycgctgtgtctctgtgagcaaggcgcaag	70
QY	61	GAATTCTATGTGAAACCTAAACAGTCTGTGAGCCCTGGAACTCCACCTCAGACAGAAATGAA	120
DB	71	gaattctgtgtgaactaaacagctgtgagcccttgaaacctcgctcagagagaatgaa	130
QY	121	GGATATCGACATAGGAAAAAGAGTATATCATCCCAAGTCTGTGGTATTTGAAAGTGTGAGGA	180
DB	131	ggatatcgacataggaanaagatatactatccaccagctcctgggtatagaagtgtgaaggga	190
QY	181	GAGAACCGACACTTCTGGGACACGCACAGAGACCGTGAAGATTCCAACTTCAGAGGAACTCG	240
DB	191	gagaaacccagcactctctgtggaacgcacaagacccgtgaagattccaaagttcagagaaactcg	250
QY	241	ACCGTTGGAAATGCCAAGATGGCTTGGAAACAGCGCCGAGCCGACGAGGCGCTCTCTTTGA	300
DB	251	accgcttggaatgaccaagatgctcttggaacaacgcagcccgacgagcgctctctcttga	310
QY	301	TGCGTCCATGCAATTCYACGTACAGAAATCCGTGGATGAGAGAGCATCCCAAGGAAAAAGTACA	360
DB	311	tgccctccatgcatctctcagcgcagaaatccgtgtgagagaatctcccaaggaaagtacaa	370
QY	361	TCATAGGCTTGAGTGTCTGTGAAGCCCATTCGCGACTTCTTGCAACACACAGACCCACGTGGA	420
DB	371	tcatagcttgtagtgcctctgaagcccatccgcgtactctccaacaacacagacccagctgga	430
QY	421	CAATCTCTGGGCTTTTCTCTGTATGAGCTTTTCTGGGCTTTCTCTCTGTGGCCGCTGAGG	480
DB	431	caatgcttgagctctctctccgtgaagaatctcttcgtgctctctctctctcgtcgtgtgac	490
QY	481	CCACAAGAGGGGGAGCTCTCTAATGAGAAAGCTGTGGTCTCTGTCCAAAGCAGCAGTCTTC	540
DB	491	ccaacaagaaggggagcgctctcaatggaagacgtgtgtgtctctgtccaagaagcagatcttc	550
QY	541	TGACGTGAATCTGCAGAGACTTGAAGAGACTGTGGCAAGAAAGCTGAATGAAGTTGGGCC	600
DB	551	tgaagtgaaactgcagaaagactagaagacgtctgtgcaagaagagcttgaaatgaactgtggcc	610
QY	601	AGACGCTCTTCCCTCGCAAGAGGTGTGTGATATCTTCTGGCGCACAGGGCTACATCCGTGC	660
DB	611	agaagctgtctccctccggaagggtgtgtgtgatactctctgcgcgaacgcagctcaatccgtc	670
QY	661	CATCTGTGCGCTGTATGATCAGCAGCACTGTGCTGAGCTTCAAGTGAACACACCTTCATGTGAA	720
DB	671	catgctgtgcctgtatgatacaagcaagcgtgcgtgtcttaagtgagaccagcctcatagtga	730
QY	721	ACACCTCTTGGAGTATACCAACGAGCAACAGAGCTTAACCTGCAGTACAGCTTGTGTTAGT	780
DB	731	acaactcttgagataataccagcaacagagctcaactctgagtaacagctgtgtgttagt	790
QY	781	GCTGGGCGCTCCCTCGACGAGAAATCGGCGGTGTGGTGGCTTGCATCGACTTGGGGATT	840
DB	791	gctgggctctccctccgaaagaaatcgtgcgtctctgtgccttgacactgcaactgtggaa	850
QY	841	GAATTACCGAAACCGGTCTCCGCTTGCGGGGGCCATCTTAACATCGCATTTAAGAAAGAT	900
DB	851	gaattaccggaacggtgctcgctctgtgcgggggccaatccataacatgcatctaagaagat	910
QY	901	CGTTAAGTTTAAAGACATTAAAGAGAAATCCCTGGGTGAGCTCATCAACTTTGCTTCCAA	960

|||||
Db 911 ccttaagtaaaacatlaaagaalccctggtgagctcatcaacatltgtccaa 970
961 CGATGGCAGAGAAATGTTTAGGAGCAGCCGTTGGCAGCTGTGGCTGGAGACCCGT 1020
|||||
Db 971 cgaatggcagaagaatgttttagagcagccggttgagcagctgtgtgtgaggaacccgt 1030
1021 TGTGGCATCTTAGCAGATGATTAATTAATGTAATTAATTTCTGGGACCAACAGCTCTGGG 1080
|||||
Db 1031 tgttgccatcttaagcatgatttataatlatcttctgagccacaacagcttccggg 1090
1081 ATGACGCTTTTATCCTCTTTTACCCAGCATGATGTTTGATCAGAGGCTCAGAGATA 1140
|||||
Db 1091 atcaagcttlttatacctcttcaaccagaaatgattgtgacatcaagctccacaagata 1150
1141 TTTTCAGAGAAAATGCTGGCCGACGAGATGACGTCGTCAGAGAAATGAAATTTCT 1200
|||||
Db 1151 ttccagagagaataatgctgtgcccagcatgaaagtgtccagaagaatgaaatgaaatctt 1210
1201 TACTTACATTAATTTATCAAAATGTATGCTGGTCAAAACATTTTCTCAGAGTGTTC 1260
|||||
Db 1211 tactacattaaatttatacaaatgtatgctgtggtcaaaacatttctcagagtgttca 1270
1261 GAAAAATCCGAGAGAGAGAGCTCGGATTTGAAAAAGCCGGGTACTTCCAGACATCAC 1320
|||||
Db 1271 aaaaatccgagagagagagcgtcagatattgaaaaagccggtacttccaggttacc 1330
1321 TGTGGGTGCTGCTCCCATTTGCTGGTGGTGGATGGCCAGGTGTCACCTTCTGTTCAAT 1380
|||||
Db 1331 tgtggt 1390
1381 GACCTTGGCTTCGATGTGACAGCAGCAGCAGCTTTCACAGTGTGACAGTCTTCAATTC 1440
|||||
Db 1391 gacctgt 1450
1441 CATGACCTTTGCTTTGAAAGTAACACCGTTTTCAGTAAAGTCCCTTCAGAGAGCTCAGT 1500
|||||
Db 1451 catgacctgt 1510
1501 GGGCTGTGACAGATTTAAAGATTTGTTTAAATGAAAGAGTTCACATGATTAAGAAACAA 1560
|||||
Db 1511 ggcgtgtgacagatttaaggttgtttcttaatgaaaggttcaactatgataaagaacaa 1570
1561 ACCAGCAGTCTCAGTCATCAAGATAGAGATGAATAATGCCACTTGGCATGGAGCTCCTC 1620
|||||
Db 1571 accagcagcttccatcaatcaagatagaaatgaaacacgttgcagtcgtgtgtgtgtgt 1630
1621 CCATCCAGTATCCAGAACTCGCCCAAGCTGAGCCCAAAATGAAAAAGACAGAGGGC 1680
|||||
Db 1631 ccactccagatccagaactcgcaccaagctgaccccaaaatgaaaaagaacaagaaggtc 1690
1681 TTTCCAGGGGCAAGAAAGAGAGGTGAGGAGCTGACAGCAGCTAGCATCAGGGGTGCT 1740
|||||
Db 1691 ttccaggggcaagaagaagaaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtg 1750
1741 GGCAGACAGAAAGGCCACTCTCTCTGACAGTGAAGAGAGGCGGCCAGTCCCAAGAGGA 1800
|||||
Db 1751 ggcagagcaagaagaagccaactcctctgtgacagtgacgagcaggtcccgagaagga 1810
1801 AGAAGGAGAGACATCCACCTGGGGCAGCTGCGCTTACAGAGGACATCGCACACATCGA 1860
|||||
Db 1811 agaaaggaagaacatccactgtgcccacgtgcttacaagaagacacgtgcaacacgaa 1870
1861 TCTGGAGATCCAAAGAGGTAAACTGTTGGAATCTGCGGACAGTGTGGAAAGTGGAAAAA 1920
|||||
Db 1871 tctgagagatccaagaggtgtaactgtgtgaaatctgtgcaaggtgtggaagtgtgaaaaac 1930
1921 CTCTCTATTTTACGCAATTTTATAGCCAGATGACGCTTCTGAGGGGACATTCGAATCAG 1980
|||||
Db 1931 ctctctattcttcagccatttaggcccagatgacgcttcttagaggtgagcagcatltgaaatcag 1990
1981 TGGAACTTGGCTTATGTGGCCGAGGCGCTGGAATCTGAATGCTCTGAGAGACAA 2040
|||||

|||||
Db 1991 tgaacctgtgcttatgtgcccagcaggtgtgatactcctaatgtctactctgaaagacaa 2050
2041 CATCCTGTTTGGGAAGAAATATGATGAAGAAAGATACAACTGTGCTGAACACCTGCTG 2100
|||||
Db 2051 catcctgttttggaaggaataatgataagaagaataacaaactctgtgtgaaacagctgtgt 2110
2101 CCTAGAGCCTTACCTGGCCATTTCTTCCAGACAGGACTGACGAGATTTGGAGAGCCAGG 2160
|||||
Db 2111 cctgaggtcctgacctgtgccaattctccagcagcagctgtgagatltggaagcagag 2170
2161 AGCCAACTGAGCGGTGGGCGAGCCAGAGAGATCAGCTTCCCGGCGCTTGTATGTA 2220
|||||
Db 2171 agccaactgtgagcgtgtgagcagcgcagagagatcaagcttgcggcgctgtatgtga 2230
2221 CAGAGCATCTACCTCTGAGAGACCCCTTCAGTGTCTTACATGCTTGGGCAACCA 2280
|||||
Db 2231 caggaagatctacatctcgtgagacccctcagctgagcttagatgtccaatgtggaacca 2290
2281 CATCTTCAATATGCTCTATCCGGAACATCTCAAGTCCAAAGACATTTCTGTTGTACCCA 2340
|||||
Db 2291 catctcaatagtgctaccggaacaactcaagttccaagaacagctgtgtgtgtaccca 2350
2341 CCAGTTACAGTACCTGGTGTGATGATGAAATGATGATCTTCATGAAGAGGCTGTATTAC 2400
|||||
Db 2351 ccagttcaagtaacctgt 2410
2401 GGAAGAGGCAACCCATGAGAGATGATGATTAATGATGATGATGATGATGATGATGATG 2460
|||||
Db 2411 ggaagagagaccccatgagagacatgaaatlttaatgtgtgtgtgtgtgtgtgtgtgtgt 2470
2461 TAACCTGTTGCTGGGAGAGACACCGCCAGTTGAGATCATTTCAAAAAAGAAACCACTGG 2520
|||||
Db 2471 taacctgt 2530
2521 TTTACAGAGAGATCAACAAGAGGAGGCTTAAACAGATCAATTAAGAGAAAAAGC 2580
|||||
Db 2531 ttccagaagaagttcaacaagaagaaggttcccaaaacagatcaagtaagaagaagaagc 2590
2581 AGTAAAGCCAGAGAGAGAGGAGGAGCTTGTGACAGCTGGAAGAGAAAGGAGGCTTACGTC 2640
|||||
Db 2591 agtaagaagcagaagaagaaggt 2650
2641 CTGCTCAGTATATGCTCTATCATCCAGGCTGCTGGGGGCGCCCTTGGCATTCCTGGTAT 2700
|||||
Db 2651 ctgtctcagatatagt 2710
2701 TATGGCCCTTTTCAATGCTGAATGTAGGAGCAGCAGCCTTTCAGACCTGTGTGAGTTA 2760
|||||
Db 2711 tatgtcccttctcagtgatgt 2770
2761 CTGGATCAAGCAAGAGAGAGGAGGAGACCACTGTGACTGTGAGGAGAGACCTGCTGAG 2820
|||||
Db 2771 ctgtatccaagaagaagcgt 2830
2821 TGAAGCATGAAGAGCAATCTCTATGCAAGTACTATGCGCATCTACGCTCTCAT 2880
|||||
Db 2831 tgaagatagaagaacatctctatagcaatgactatgcaagcactcaagccttccat 2890
2881 GGCAGTATGCTGATCTGGAAGGACATTCGAGAGTGTCTTGTCAAGGAGCAGCTGCG 2940
|||||
Db 2891 ggcagatcatgt 2950
2941 AGCTTCTCCCGGCTGCAATGACAGGCTTTTCCGAGAGATCTCTTGAAGCCCTATGAAGTT 3000
|||||
Db 2951 agcttctcccggtctgatatgagcagcttltccgaagatccttccgaagccctatgaaglt 3010
3001 TTTTGAACAGACCCCAAGAGAGATTCACAGGTTTTCACAAAGATGATGATGATGATGAT 3060
|||||
Db 3011 ttctgaaagagacccccaagagagatltcaaaaglttccaagaacatgagtgagt 3070
3061 TGAGCTGCGGCTGCTTTCAGGCGAGAGATTTATCCAGAACGTTTATCTGTGTTCTT 3120
|||||
Db 3071 tgaagtgaggt 3130

3121 CTGTTGGGAATGATCGAGAGAGTCTCCGTTCTTGTGGCAGTGGGCCCCCTTGT 3180
3131 ctggtggaatgatacgcagagagctcccggtctccttgcgtgcagtgggcccccttgc 3190
3181 CATCCTCTTTTCACTGCTGACATATGTCACAGGGTCTGATTCCGGAGCTGAACGCTCT 3240
3191 catcctctttcagtcctgcacatgtgtccaggttctgtatctggagctgaaagcgtct 3250
3241 GGACAAATATCAGCAGTACCTTTCTCTCCCAATCAGCTCAGCATACAGGCGCTTGC 3300
3251 ggacaatacagcagctaccccttccctccacacacagctccagcatacagggccttcg 3310
3301 CACCATCAGCGCTACATATAAGGGCAGAGTTTCTGCANATACAGAGACTCTCTGA 3360
3311 caccatcagcgctacataaaggcagaggttctgcacagatacagagcgtctgga 3370
3361 TGACAAACCAACTCCTTTTCTTTTGTACGTTGCGATGCGGTGGCTGGCTGGCT 3420
3371 tgacaaccaagctcctcttcttcttcttcttcttcttcttcttcttcttcttcttct 3430
3421 GGACCTCATCAGCATGCGCTTCATCAGCAGCAGGGGCTGATGATGTTTATGCACGG 3480
3431 ggacctcatcagcatgcctcatcacacacacagggcgtgatgtatcgttcttcttcttct 3490
3481 GCAGATTTCCCCCAGCGCTATGGGGGTCTGCGCATCTCTATGCTGTCCAGTTAACGGGCT 3540
3491 gcagattccccagcctatgcgggtctgcacatctcttcttcttcttcttcttcttcttct 3550
3541 GTTCAGATTACGGTCAGACAGTGCATGAGCAGAGAGCTCATCACCCTCGTGGAGAG 3600
3551 gtccaagttaacggtcagacacgtgccttgagacagagctcgtatcaccctcgttgagag 3610
3601 GATCATCATCATATTAAGACTGTCTGTGGAAGCACCCTGCCAGAAATTAAAGACAAGC 3660
3611 gatcatcatcatcatataagactctgtccgtgaagcaccctgcagaattaaagaacagc 3670
3661 TCCCTCCCTCAGTACTGGCCCGAGGAGGAGGTGACCTTTGAGAACGCAATAGATAGTA 3720
3671 tccctccctcagctgagcccgagagaggtgagcttgcagcagcagagataagagta 3730
3721 CCGAGAAAACCTCCCTCTGCTCTTAAGAAAGTATCTTCACAGATCAAACTTAAGAGAA 3780
3731 ccgagaaaaacctccctctgtctctaaagaagtatccttcaagatcaaaacctaaagaga 3790
3781 GATTGGCATTTGTGGGGCGAGATCAAGGAATCTCGCTGGGGAGTGGCCCTCTTCGG 3840
3791 gattggcatltgaggcgagacagatcaggaagtcctcgtgggagtggccctcttcgg 3850
3841 TCTGGTGAAGTTATCTGAGGTGATGATCAAGATTTGATGAGAGTGAATAGATATTGG 3900
3851 tctggtagtatactgaggtcgtacaaagatcgtagtagagaaatcagatgatalttgg 3910
3901 CCTTCCCACTCCGAAACAACCTCTATCATCTCAATCTCTCAAGACGGTCTTTAGTGG 3960
3911 ccttcccaacctccgaagaacactctctatcatcttctcaagggccgggtgttcaagtg 3970
3961 CACTGTCAAGATCAAAATTTGGACCCCTTCAACCACTGAAAGCAGATTTGGGATGC 4020
3971 cactgtcagatcaaaatttggacccctcaaccagtaactgaaagcagatttggatgc 4030
4021 CCTGAGAGGACACATGAAGAAATGATTGTCTAGCTACTCTGAACCTTGAAATCTGA 4080
4031 cctgagagggacacacatgaagaagtattgtctcagctacacctcgaaccttgaacctga 4090
4081 AGTGATGAGAAATGGAGTAACCTTCTAGTGGGAGACGCGCTTTGTGATAGCTAG 4140
4091 agtgatgagaaatggagtaaccttctcagtgagggaacgagcctctgtgtatagctag 4150
4141 AGCCTGCTCGCACCTGTAAGATTTGATTTATGATGAAGCACAAGCTGCATGAGCAC 4200
4151 agcctgtctcgccactgtaagatctgtatttagatgaagcacaagcagctgcatagcac 4210

4201 AGAGACAGACTTATTGATTCAAGAGACCATCCAGAACATTTGACAGTTCACATGCT 4260
4211 agagacagacttattgattcaagaagacacatccgagaagatttgcagactgtacatgct 4270
4261 GACCAATGGCCATCCCTGACACAGGTTCTAGGCTCCGATAGATTAAGTGTGGCCCA 4320
4271 gaccatggccatccctgacacaggttctagagctcccgtaagatattgtgtgcccac 4330
4321 GGGACAGGAGGAGTGTGAGACACCCCATCGCTCTTGTGTCCAAACGACATGCCGAT 4380
4331 gggacagaggtgtgaggttgcagaccacatcgctccttctgtccaaacgaagttccgatt 4390
4381 CTATGCATGTTGTGCTGTCAGAGAACAGTCCGTGCAAGGCGTCACTCCCTGT 4440
4391 ctatgcattgttgcctgtgcagagaacaagtcgtctcgaaggtctgactcctccgt 4450
4441 TGACGAAGTCTCTTTTCTTTAGAGCATTTGCCATTCCTGCTGGGGCGGGCCCCCTTATC 4500
4451 tgacgaagctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 4509
4501 GCGTCTCTACCGAAACCTTGCCTTTCGATTTATCTTTGCGACAGCATTCGGAT 4560
4510 gcgtctctacacgaaccttgccttctcgaatttatcttctgcacagcagcttcggat 4569
4561 TGGCTGTGTGTTCACTTTTAGGAGAGTCATATTGATTGATTATTTCATAT 4620
4570 tggctgtgtgttcacttttagggagagatcattgtatcattgtatcattcattc 4629
4621 TCATGTAACAAATTTAGTTTGTCTTAAATGCACTCTAAAGGTTCAAGGAACCGT 4680
4630 tcattgaacaaatttagtttcttcttcttcttcttcttcttcttcttcttcttcttctt 4689
4681 TATTTAATTTATCAGAGCGCTATATATACGTTTATACGTTATCTATATAT 4740
4690 tatttaatttatacagagccataatgaagcttatacgtgtgactatataata 4749
4741 AATTGTCATAGCCTATATTATACAGTAAGTAAGTAAAGCTGTTATTTATTTAAAT 4800
4750 aattctgatacagccataatcagtgaaagtgaagctgttatttataataat 4809
4801 AAGCACTGTGCTAAAAAA 4819
4810 aagcactgtgctaataaca 4828

RESULT 4
AA294745
ID AA294745 standard; cDNA; 5838 BP.
XX
AC AA294745;
XX
DT 01-AUG-2000 (first entry)
XX
DE Human ATP binding cassette ABC5 (MRP5) cDNA.
XX
KM ABC5; ATP binding cassette; human; cholesterol; lipid disorder;
KM atherosclerosis; lipid disorder; dyslipidemia; psoriasis;
KM lupus erythematosus; diagnosis; gene therapy; MRP5;
KM multidrug resistance associated protein; chromosome 3q25-27;
KM Dublin Johnson syndrome; hyperbilirubinemia; ss.
XX
OS Homo sapiens.
XX
PN WO200018912-A2.
XX
PD 06-APR-2000.
XX
PF 21-SEP-1999; 99WO-EP06991.
XX
PR 25-SEP-1998; 98US-0101706.
XX
PA (FARB) BAYER AG.

PI Schmitz G, Klucken J;
XX WPI; 2000-293151/25.
XX Adenosine triphosphate binding proteins useful for identifying agents
PT for treating atherosclerosis and other inflammatory disorders -
XX
XX
PS Claim 9; Page 123-125; 154pp; English.
XX
XX The present sequence is that of human ATP binding cassette
CC subfamily C protein ABC5 cDNA. The cDNA was identified using a
CC differential display method in which monocytes from peripheral
CC blood were subjected to macrophage differentiation and cholesterol
CC loading with acetylated low density lipoproteins and subsequent
CC degrading with high density lipoprotein (HDL3) to identify
CC cholesterol sensitive genes. The gene maps to chromosome 3q25-27
CC and is also termed MRP5 (multidrug resistance associated protein).
CC This is the gene locus for Dublin-Johnson syndrome, a disorder
CC associated with mild chronic conjugated hyperbilirubinemia.
CC The invention provides cholesterol-sensitive ABC genes (see
CC AB294734-63). These genes, and polypeptides encoded by them,
CC can be used for diagnostic and therapeutic applications, and for
CC biochemical or cell-based assays to screen for pharmacologically
CC active modulator compounds useful for the treatment of lipid
CC disorders, atherosclerosis or other inflammatory diseases such as
CC psoriasis and lupus erythematosus.
SI Sequence 5838 BP; 1422 A; 1432 C; 1509 G; 1475 T; 0 other;

Query Match 98.98; Score 4792.6; DB 21; Length 5838;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 4809; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 1 GGGTCATGCTCGGGAGGCTGGTGGAGGGGCGGCGGTTGCTGGAGCAGGGGCGCAG 60
DB 11 ggcctacgctcggagagctggtgagcggcgctgctgctgagcgggag 70
QY 61 GAATTCGATGTGAATACAGTCTGTGAGCCTCGAACCCTCAGTCAGAGATGAA 120
DB 71 gaattcgaatgtgaatacagctgtgtgagcctcggaacctccgccaagagaagaa 130
QY 121 GGATATGACATAGGAAAGAGTATATCATCCCAAGTCTGGTATAGAGTGAAGGA 180
DB 131 ggaatcgaatgtgaatacagctgtgtgagcctcggaacctccgccaagagaagaa 190
QY 181 GAGAGCAGACCTCTGCGAGCGCAGACAGACCGGAGATTCAGATTCAGAGAACGCG 240
DB 191 gagaacagacactctcggagcagcagagacgtgaaatccaaatccaggaagaaactcg 250
QY 241 ACCGTTGGAATGCCAAGATGCTTGAAGACAGCCGAGCGAGGCGCTCTCTTGA 300
DB 251 accgttggaatgccaaagatgcttgaagacagcccgagccgagggcctctcttga 310
QY 301 TGCCTCATGCAATCTCAGATCCTGAGATGAGAGCAGATCCCAAGGAAATACCA 360
DB 311 tgcctcatgcaatctcagatcctgagatgagagcagatcccaaggaagaaatacca 370
QY 361 TCATGAGTGTAGTCTGGAAGCCCATCCGACACTACTTGGCAACACAGCCAGTGA 420
DB 371 tcatgagtgtagtgcttgaagcccatccgactacttccaaacacagccccaagtga 430
QY 421 CAATGCGGGCTTTTTCGATGATGATCTTTCGTGGCTTCTCTGCGCCGTGAGGC 480
DB 431 caatgctgggcttttccgatagaacttttcgtgcttctctcctcgcgcgtgagcc 490
QY 481 CCACAGAGAGGGAGCTCTCAATGAGAGAGTGTGTTCTCTCAAGACGAGTCTTC 540
DB 491 ccacagagagggagctctcaatgagagagctgtgtctctcctcgaagcagctcttc 550
QY 541 TGACGTAACTGCGAAGACTAGAGACATGTGGCAAGAGAGCTGAATGAATGGGCC 600
DB 551 tgaactgaactgacgaagactagagacatgtggcgaagagagctgaatgaatgggcc 610

QY 601 AGAGCTGCTCCCTGCGAAGGTTGTGTGATCTTTCGCCAGCAGGCTCATCTGTC 660
DB 611 agagctgctccctcgcgaaggtgtgtgatacttctccgcagccagcctcctcgtc 670
QY 661 CATGCTGCTGATGATACAGCAGCTGCTGCTTCACTGAGACAGCTTCAATGAGTGA 720
DB 671 catgctgctgatagatacagcagctgctgcttcaactgagacacccctcattgtaa 730
QY 721 ACACCTCTGAGATATACAGCAGCAGAGCTAATCCAGTACAGCTGTTGTAGT 780
DB 731 acacctctgagatatacagcagagctcctcaactgagctgagctgctgagct 790
QY 781 GCTGGGCTCCTCTGACGAGAAATGTCGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 791 gctgggctcctctcgcagagaaatcgtgctgctgctgctgctgctgctgctgct 850
QY 841 GAATTCGCAACCGGTGCTGCGGAGGCGCATCTTAACATGAGATTAAGAAAT 900
DB 851 gaattcgcgaacccggtgctgctgctgctgctgctgctgctgctgctgctgct 910
QY 901 CCTTAAGTAAAGAACATTAAGAGAAATCCCTGGTGAAGCTCAATCTGCTCAA 960
DB 911 ccttaagttaagaaacatataagagaaatccctgggtgagctcattcattcctcaa 970
QY 961 CGATGGCAGAGAAATGTTGAGCAGACGCGTTGGACGCTGCTGCTGCTGCTGCTGCT 1020
DB 971 cgatggcagagaaatgttgagcagacgctgctgctgctgctgctgctgctgctgct 1030
QY 1021 TGTGTCATCTTAGCATGATTTATATGTAATTTATGTTGGGACCAACAGCTCTGAGG 1080
DB 1031 tgtgctcatcttagcatgatttataatgtaatttatttggaccaaagctctcgtg 1090
QY 1081 ATCAGCTGTTTATACCTCTTACCCAGATGATGTTTCATCAGCGCTCACAGCATA 1140
DB 1091 atcagctgttttattacctcttaccagcaatgattgttgcatacagctcacaagcta 1150
QY 1141 TTTGAGGAAATGCGTGGCGCCGACGAGTAAACGTGTCCAGAAATGAATGAATGTTCT 1200
DB 1151 tttagagaaatgctgctggcgccagagatgaacgtgtccagaagaatgaatgattct 1210
QY 1201 TACTTCACTTAATTTATCAAAATGATGCTGGGTCAACATTTTCTGAGAGTGTCA 1260
DB 1211 tacttcaacttaatttatacaaaatgatactgctggtcaaacatttctcagaatgttca 1270
QY 1261 GAAATCCGCGAGGAGGAGCGCTGGATTTGAAAGAACCCGGTACTCCAGAGCATCAC 1320
DB 1271 aaaaatccgcgagagagagctgctggaatatgtgaaagacgggtactcagggatcac 1330
QY 1321 TGTGGGTGTGCTCCCATTTGATGTGATTTGCCAGCTGTGACCTTCTGTTCAATAT 1380
DB 1331 tgtgggtgtgctcccatTTGATGTGATTTGCCAGCTGTGACCTTCTGTTCAATAT 1390
QY 1381 GACCCGTGGGCTTCGATCTGACAGCAGCAGAGCTTTCACAGTGTGACAGCTTCAATTC 1440
DB 1391 gacctgtgcttcgactcgcagcagcagcagcttccacagtggtgaaagcttcaatttc 1450
QY 1441 CATGACTTGTGCTTGAAGTAAACGCTTTCAGTAAAGCCCTCGAAGAGCTCAGT 1500
DB 1451 catgacttgtgcttgaagtaaacgcttctcagtaagccctcctaagaagctcagtc 1510
QY 1501 GCGTGTTCAGATTTAAGATTTGTTTCTTAATGAGAGAGTTCACATGATTAAGAACAA 1560
DB 1511 gcggtgttcagatTTAAGATTTGTTTCTTAATGAGAGAGTTCACATGATTAAGAACAA 1570
QY 1561 ACCAGCCAGTCTCAATCAAGATAGAGATGAAGAAATGCCACTTGGAGTGGAGCTCTC 1620
DB 1571 accagccagctctcaatcaagatagagatgaagaaatgccacttggagtggagctctc 1630
QY 1621 CCAATCCAGTATCCAGAACTGCCCCAAGCTGACCCCAAAATGAAAAAGCAGAGGCGC 1680
DB 1631 ccaatccagtatccagaactgcccccaagctgaccccaaaaatgaaaaagcagaagggc 1690


```

1681 TTCCAGGGGCAAGAAAGAGTGTAGGCGAGCTGCGAGCGACTGTGACATCAGCGGTGCT 1740
1691 ttccaggggcaagaaagagtgatgagcagctgcagcgcactgagcatcagggcggctgtc 1750
1741 GGCAGAGCAAGAAAGGCCACCTCTCTCTGAGCAGTGAAGCGGGCCAGTCCGGAAGAGA 1800
1751 ggcagagcaagaaagccacctctctctcgtgacagtgacgagcgccagcccgaaagaga 1810
1801 AGAAGGCAAGCAGATCCACCTGGGGCACCCTTGCGCTTACAGAGACACACACATGCA 1860
1811 agaagggcaagcagatccaccctgggaccctgcgtctacagaggaaacatgcacagatcga 1870
1861 TCTGGAGATCCCAAGAGGGTAAACTGTGTGSAATCTGCGGCACTGTGGAAAGTGAAGAAAC 1920
1871 tctggagatcccaagagggttaaactigtgtgaaatctgcgscagtgcygggaagtggaanaac 1930
1921 CTCTCTCATTTTCAGCCCATTTTAAAGCCAGATGACGCTTCTAGAGGGCAGCATTTGCAATCAG 1980
1931 ctctctcatlttaagccatttttaggcagatgagcgtctcttaaggggcagcatltgcgaatcag 1990
1981 TGGAGCCTTCGCTTATGTGGGCCACAGCGCTCGATCCATGCTACTCTGAGAGACAA 2040
1991 tggaaaccttcgcttattgycgcccagcagcgtgatacctcaatgctactctgaagagacaa 2050
2041 CATCCTGTTTGGGAAGAAATATGATGAAGAAAGATACAACTGTCGTGAACAGCTGCTG 2100
2051 catcctgtttgggaagaaatatatgaaagaaagatacaactctgtcgtgaacagctgctgt 2110
2101 CCTGAGGCGCTGACCTGGCCATTTCTTCCAGCAGCAGCTGACGAGATTTGAGAGCGAGG 2160
2111 cctgagcgctgacctggccatttctccagcagcagcctgacgagatltgggagcgaggt 2170
2161 AGCCAACTGAGCGGTGGGCGACGCCAAGATCAGCCTTGCCCGGGCTTGTATATGTA 2220
2171 agccaaactgagcggtggcgacgcgcagaaagatcagccttgcgcgggccttataagtga 2230
2221 CAGAGCATCTACATCTCGAGAGACCCCCCTCAGTGCCTTAGTGGCCATGTGGGAACCA 2280
2231 cagagcatctacatctccgagagacccccctcagtgcttagagtgcccatgtggtgaacca 2290
2281 CATCTTCATTAAGTGCTATCCGGAACATCTCAAGTCCAAAGACAGTTCGTTTGTAAACCA 2340
2291 catcttcaatagtgctatccggaacatctcaagtcgaagacgttctgtgttaacca 2350
2341 CCAGTTAAGATACCTCGGTGAGCTGTGATGAGTGAATCTTACGAAGAGCGCTGATTAAC 2400
2351 ccagtttaagatacctcggtgactgtgatagtgaatgtcttcaatgaagagggctgtataaac 2410
2401 GGAAGAGGCAACCCATGAGGAACGTATTAATGGTGAATGACTATGCTACCATTTTAA 2460
2411 ggaagaggcaacccatgaggaacgtatgaatgaatgaatgaatgaatgaatgaatgaatgaat 2470
2461 TAACTGTTGCTGGGAGAGACACCGCCAGTTGAGATCAATTCAAAAAGGAACACAGTGG 2520
2471 taactgttgcctgggagagacacccgcaagtgtagatcaatctcaaaaaaggaaccaagtggt 2530
2521 TTCACAGAAGAGTCAACAAGACAGGCTCTTAAACAGGATCAATTAAGAGGAAAGAAC 2580
2531 ttcaacaagaagagtcacaagaagaagtgctctaacaacagatcagtaagaaggaanaaac 2590
2581 AGTAAGGCAGAGGAAGGGCAGCTGTGACGTGAGGAAGAAAGGCGAGGCTTCAGTGC 2640
2591 agtaaggcagaggaagggcagctgtgacgtgaggaagaaaggaaggaagggcttcagtgcc 2650
2641 CTGCTCATATATGATGTGCTACATTCAGGCTGCTGGGGCCCTTGATTCCTGTTAT 2700
2651 ctgctcatatgatgtgctacatcagctgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgt 2710
2701 TATGGCCCTTTTCAATGCTGAATGTAGGACAGACCGCCTTCACAGCCTGGTGGTTAGTTA 2760
2711 tatggcccttttcaatgctgaatgttaggacagacccgcttcacagcactggtggttgatga 2770
2761 CTGGATCAAGCAAGGAGCGGGAACACCACTGTGACTGAGGGAACGAGACTCGGTGAG 2820
2771 ctggatcaagcaagggagcggggaacaccactgtgactgaggggaacgagactcgtgtgag 2830
2821 TGACAGCATGAAGACACATCTCTATATGACGACTACTATGCGACATCTACGCCCTTCAT 2880
2831 tgacagcatgaagacacatctctcataatgacgtactatgcagcatctaacgctctccat 2890
2881 GGCAGTCATGCTGATCCGTAAGGCATTCGAGAGTGTGCTTGTCAAGGGCAGCGTGGC 2940
2891 ggcagtcagtcgatacccgaaagccatcagagatgtgtcttgcgaagggcagcgctgcg 2950
2941 AGCTTCCTCCCGGCTGACATGACAGAGCTTTTCCGAAGGATCTTCCAAAGCCCTATGAAGT 3000
2951 agcttcctcccggtcgtacatgacgagcctttccgaaagatcccttcgaagccctatgaagt 3010
3001 TTTTACACAGACCCCCACAGGGAGATTCTCAACAGGTTTTTCCAAAAGCATGTGATGAAGT 3060
3011 ttttacacagaccccccacagggagatctcaacaggttttccaaaagcatgatatgaagt 3070
3061 TGACGTMGGGCGTCCGCTTCAGCGCGAGATGTTCAATCCAGAGGTTATCCGTGTCCT 3120
3071 tgaagtgcggtcgccgtctccagggccgaagatgttcaaccgaacgtataccgtgttctc 3130
3121 CTGTGTGGGAATGATCGAGAGATCTTCCCGTGTCTCTTGTGGCAGTGGGCGCCCTTGT 3180
3131 ctgtgtgggaatgatcgagagatcttcccggtgtctcttctgttcctgtggaagtgggccccctgt 3190
3181 CATCTCTTTTTCAGTCTGACATCTGTCACATGTTCTCCAGGGCTCTGATTCGGAAGCTGCT 3240
3191 catctctttttagcttcgtcacatctgtctccaggttccaggttccggaagctgaagcgtct 3250
3241 GGACAAATATCAGCGAGTACACTTCTCTCTCCACATCAGCTCCACATTCAGAGGCTTGC 3300
3251 ggacaaatatacagcgagtcacacttctctccacatcaagtcacagatacagggccttgc 3310
3301 CACCATCCACGCGCTCAATTAAGAGGCGAGAGTTCGTGCACAGATATCCAGAGACTCTGGA 3360
3311 caccatccacgctacataaataaaggcagaggttctgtcacagaataccagagctgtgtga 3370
3361 TGACAAACCAAGCTCCTTTTGTGTTTACGTGTGCGATGCGGTGCGTGGCGGCT 3420
3371 tgacaaccaagctccttttcttctgtttagtgtgtgagtggtgtgcgtgcgtgcgtgc 3430
3421 GGACCTCATACAGATCGCCCTCATCATCACACACGAGGGGCTGAGTCTGTTATATGACAGG 3480
3431 ggacctcatcaagcatcgccccctcatcaaccaaccaagggtcgtgatacgtcttatagcag 3490
3481 GCAGATTCGCCAGCGCTATGCGGGTCTGCGCATCTCTTATGCTGTGTCAGATTAAAGGGGCT 3540
3491 gcaagatcccccaagcctatgtcggtgtcgtgcacatcccttattgtgtccagttaaaggggt 3550
3541 GTTCCAGTTTACGGTCAGACTGCGATCTGAGACAGAAAGCTCGATTCACCTCGGTGGAGAG 3600
3551 gtccaagtttaagtcagagctgtgcatctgagacagaaagcttcgaacctcgtgtgagag 3610
3601 GATCATCATCTACATTAAGACTGTCTCTTGAAGCACCTGCGCAAAATTAAGAAAGGC 3660
3611 gatcatcatctacatlaagacttctcttggaaagcaccttcgcagaatlaagaaagagc 3670
3661 TCCCTCCCTCGACTGCGGCCACAGAGGAGAGTGAACCTTTGAACGCGAGATGGGTA 3720
3671 tccctccctcgactgcccccaaggaaggaagtgacaccttggaaagcagagataagtgata 3730
3721 CCGAGAAAACCTCCCTCTGTCGTTAAAGAAATATCTTCACAGATCAAACTTAAGAGAA 3780
3731 ccgagaaaaacctccctctgtctcttaagaagtatcccttcacagatcaaaccttaagagaa 3790
3781 GATTGGCATTTGTGGGGCGGACAGGATCAAGGAAAGTCTGCTGGGGATGAGGCCCTTCTCG 3840
3791 gatggcatlttggtggggcgagacagatcaaggaagtgacgtcctgcgtgggataggcccccttc 3850
3841 TCTGTGAGTATCTGAGGCTGCATCAAGATTGATGGAGGAGAAATCAGATATTGG 3900

```


[illegible]

QY	1387	GGGCTTGATCTGCACGACACACAGGCTTTCACAGTGGTGACAGCTTTCATTCGATGAC	1446
Db	1321	gggtctcgatctgcacgacgacacaggtcttcacgtgtagcagcttccaattccattgac	1380
QY	1447	TTTTGGCTTGAAGTATACACGGTTTTTCAGTAAAGTCCCTTCAGAAAGCTCAGTGGCGT	1506
Db	1381	tttgccttgaagtaaacaccggtttccagaaagtcccttcagaagctccagtgccgt	1440
QY	1507	TGACAGATTTAAGAGTTTGTTTCTAATGAGANAGGTTCACATGATTAAGAACAAACGAC	1566
Db	1441	tgacagatttaaggtgttcttctaattgaaagagtctccatgatataaagaacaaaccagc	1500
QY	1567	CAGTCTTCACATCAGATAGAGATGAAAAATGCCACTTTGGCATGGACTCCTCCACTC	1626
Db	1501	cagttcctacatcaagatagagatgaaaaatgcccacttgcacatgtagctcctccactc	1560
QY	1627	CAGTATCCAGAACCTGGCCCAAGCTGCAGCCCAAAATGAAAAAAGCAAGAGGGCTTCAG	1686
Db	1561	cagtatccagaacacgcgcccaagctgcagcccccaaaagaaaaagacaagggcttccag	1620
QY	1687	GGGCAAGAAAGAGAGGTGAGGCGATGCAGCGCATCTGACATCGAGGCGTGTGGCAGA	1746
Db	1621	gggcaagaagaagagaggttgagtcagctgcagcgcactgcagcatcaagcggtgttcgaga	1680
QY	1747	GCAGAAAGGCCACTCTCCTCTGACAGATGACAGCGGCCCACTCCGAAAGGAAGAAG	1806
Db	1681	gcagaagggccaactctctctctgcagtgtagcgagcgccagctccgaagaggaagaag	1740
QY	1807	CAGCACAATCCACTGGGCGCACTGGCGTTACGAGAGAACGACGACAGTGCAGTCTGGA	1866
Db	1741	caagcaatccaactgcgccaactgcgcttaacagagaaacgcgcacgcatcgactgga	1800
QY	1867	GATCCAAAGAGGGTAAACTGGTGTGGAATCTGCGCACTGTGGGAACTGAAAAAACCCTCT	1926
Db	1801	gatccaagaggttaactggttggaaatctgcgcagtgtagggaaagtggaaaaaccctcct	1860
QY	1927	CATTTCACGCATTTTAGCGCAGATGACGCTTCTAGAGGCGACGACTTGCATCATGTGGAC	1986
Db	1861	catttcagccattttaggccaagatgacgtcttcaagggcgacatgtcaatcaatgtagaac	1920
QY	1987	CTTGGCTTATGTGGCCGACGAGCGCTGGATCCCATGCTACTGCTGAGAGCAACATCCT	2046
Db	1921	cttcgcttaatggtgcccagcagaagccggaatccccaatgctacatctgcgaagaaacatct	1980
QY	2047	GTTTGGGAGAGAAATATATATANAAGAGATACAACTTGTGCTGAACAGCTGCTGCCATG	2106
Db	1981	gtttgggaagaataatataigatgaagaagaatatacaactgtgcgtgaacagctgctgcag	2040
QY	2107	GCTGCACCTGGCCATTTCTTCCACGACGACGACCTGCAGAGATTTGAGAGCGAGGCCAA	2166
Db	2041	gctgcacctggccattcttcccgacgagcgtgcagagatttggagagcgagggccaa	2100
QY	2167	CCTGAGCGGTGGGACGGCCAGAGAGATCAGCCTTGCCGGGCTGTATAGACAGAGAG	2226
Db	2101	ccctgagcggtgggacgagccagaggaatcaagccttgcggcggtctgtatagtgaaagag	2160
QY	2227	CATCTACATCCCTGGACACACCCCTCAGGTGCTTATAGGCCATGTGGGCAACACATCTT	2286
Db	2161	catctacatctctgagcaacccctcagtgctttagtgcctatgctgcctatgtaggaacacatct	2220
QY	2287	CAATAGTGTCTATCCGGAACATCTCAAGTCCAGACAGAGTTCTGTATTACCCACAGTT	2346
Db	2221	caatagtgtctatccggaacatctcaagatccaaagacagttctgtttgtaaccacagtt	2280
QY	2347	ACAGTACTGGTTACTGCTGATGATGAATGATCTTCATGAAAGAGGCGTGTATTACGGAAG	2406
Db	2281	acagtaacctgtgtactgtatgataagtgcattctcatgaaagagggctgtatatacgyaag	2340
QY	2407	AGGCACCCATGAGGAAGACTGATGGAATTTAAATGGTGACTTGTACCATTTTAAATAACT	2466
Db	2341	agggacccaatgaagaaactgataatgaaatggtgactatgctaccatcttctaataacct	2400
QY	2467	GTTGCTGGAGAGACACCGCGACTGTGAGTCAATTCAAAAAAGAGAAACGATGGTTTCA	2526

```
||||| 2401 gtgtctggagagacaccgacagttagatcaatcacaagaagaaacagctgtgtccaca 2460
||||| 2527 GAAGAAAGTCACAAAGACAGGCTCTAAACAGAGATCAATTAAGAAAGAAAAGCAGTAAA 2586
||||| 2461 gaagaatgtaacaagaacaaaggttcttaaaacagatcaataaagaagaaagaacagttcaa 2520
||||| 2587 GCCAGAGAGAAAGGACGCTTGTGACGTGGAAGAGAAAGGAGGAGGCTTGACGTGCCCTGCTC 2646
||||| 2521 gccagagaagaagcgacgtctgtgcagcttggaagaagcgaggtcttcagtcgcccgtgtc 2580
||||| 2647 AGTATATGTTCTTACATCCAGGCTGCTGGGGGCCCTTGCGATTCCTGGTTATATAGC 2706
||||| 2581 agtatagtgtgttacaaccagctgtcgtggtggcccttgatccctgttataatagtc 2640
||||| 2707 CCTTTTCATGCTGATGATGAGCAGACCCGCTTACAGCACTGATGATGATGATGATGAT 2766
||||| 2641 cctttcattgtctgaatgtatgagcagacccgcttcagaccctgtgtgtgttactgtgat 2700
||||| 2767 CAAGCAAGAAAGCGGGAACACCACTGTGACTGAGGGAAGAGACGTCGGTGAAGTACAG 2826
||||| 2701 caagcaagaagcgaggaacacactgtgactcgagggagagactcgtgtgagtgcag 2760
||||| 2827 CATGAGAGACAAATCTCTATATGACGATCTATGCAAGATCTACGCCCTTCCATNGGACGT 2886
||||| 2761 catgaaggaacacccctcaatgacgactatgcaagacatcagccctccatagtcagtc 2820
||||| 2887 CATCTGTATCTGAAAGGCATTCGAGAGAGTGTGCTTTGTCAAAGGACGCTGCGACGCTTC 2946
||||| 2821 catgtgtactctgaaagcactatcgagaggtgtctcttcttcaagggcaagcgtcgagcttc 2880
||||| 2947 CTCGCCGCTGCATGACGAGCTTTTCCGAAAGATCTCTTGAAAGCCCTATGAAGTTTATGA 3006
||||| 2881 ctcccggtctgactagcagagcttctccgaagatcccttcgaagccctatgaagatttcttga 2940
||||| 3007 CACGACCCCGCAAGGAGGAGATTCACAAGGTTTCCAAAACATGATGATGATGATGAGT 3066
||||| 2941 cagaccccccaaggaagatctccaacagttctccaagaacagagatcgaagtcagtcagtc 3000
||||| 3067 GCGGCTGCCGTTCCAGGCGCGAGATGTTTCATCCAGAACTTTTCTGATGATGATGATGATGAT 3126
||||| 3001 ggcgctgcgcttccagcgagatgttcatcagaagcttaccgtgtgtcttctgtgtgt 3060
||||| 3127 GGGAAATGATCCAGAGAGCTTCCCGTGTCTCTGTGTGACATGAGGGGCCCTTGTATCATCT 3186
||||| 3061 gggaaatgtatcgaagagcttcccggtgtcttctgtgtgcaagtcggggccctgttcatcct 3120
||||| 3187 CTTTTCATGCTCTGCATGTTCTCCAGGAGCTGCTGATTTGGGAGCTGAAGCGCTGACAA 3246
||||| 3121 ctttcagctctgcaaatgtctctcaaggtctcagatcccgagagctgaaagcgtctggacaa 3180
||||| 3247 TATGACGAGTACACTTTCCTCTCCACATCAGTCCAGATACAGGGCCTTGCACACAT 3306
||||| 3181 tatcagcagtcacacttctctctccacatacgttccagcatalcagggccctgtccacat 3240
||||| 3307 CCAGGCGCTACAAATAAGGGCAGAGATTCTCTGCACAGATACCAAGAGCTGCTGAGATGA 3366
||||| 3241 ccaagctctacaataaagggcagaggttctctgcacagatacaagagcgtcgtgaagacaa 3300
||||| 3367 CCAAGCTCTCTTTTGTGTTAGTGTGATGATGAGGAGGCTGAGGCTGAGGCTGAGGCTGAG 3426
||||| 3301 ccaagctctcttcttcttcttcaagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 3360
||||| 3427 CATGACGATCGCCTTCATACACACACAGGGGCTGATGATGATGATGATGATGATGATGAT 3486
||||| 3361 catcagcagtcgcccctcatcacacacacagggcgtgacgtcttcttcatgcaagggcagat 3420
||||| 3487 TCCCCAGAGCTATGCGGGTCTCGGCATCTCTTATGCTGTCAAGTTAAAGGGGCTTTC 3546
||||| 3421 tccccagagctatgctgggtctcgccatcttctatgctgttcaagtgaaagggcgtgttcca 3480
||||| 3547 GTTATAGGTCAGACGAGGATGAGAGCAAGAGCTCGATTCACCTGGGAGGAGATCAA 3606
|||||
```

```
D 3481 gtttaaggtcagactgtgcatctgagacagaagctcgatctacccctcggtgtgagagatcaa 3540
Q 3607 TCACATCAATTAAGACTGTCTCTTGGAAAGCACTGCCAATTAAGAAACAAGGCTCCCTC 3666
D 3541 tcaactacaatgaactctgtctcttggaaagcaactcgtccagaattaaagaacaagcttccctc 3600
Q 3667 CCCTGACTGCCCCCAGAGGAGAGAGTACCTTTGAGAACCAAGATGAGAGTACCGAGA 3726
D 3601 ccttgcactgtgccccagagggagaggtgacatttgagaaacgcaagatgagaggtacagaa 3660
Q 3727 AAACCTCCCTCGTCCGCTAAAGAAAGTATCCCTTCACAGATCAACCTTAAGAGAGATG 3786
D 3661 aaactccctctgtctcctaagaagatcccttcaagatcaactaaagagaagatttgg 3720
Q 3787 CATGTGGGGGGAAGAGATCAGGGAAGTCTCGCTGGGATGGGCCCTTTCCTGTGCT 3846
D 3721 catgtgtggcgagacaggaatacaggaagctcctgtggtggatggcccttcgtgtgt 3780
Q 3847 GGAATTATCTGAGGCTTCATCAAGATTGATGAGTGAATCACTGATATTGGCTTTC 3906
D 3781 ggaattatctgagagctgtgcatcaagatgatatgagtgagaaatcagttatgtgccttgc 3840
Q 3907 CGACCTCCGAAGCAAACTCTATCATATCTCTCAAGAGCGCGGTGCTGATGAGTACGACTGT 3966
D 3841 cgaactccgaagcaaacctctcatcatcctccaagagcgggtgtgttcagtgagctgt 3900
Q 3967 CAGATCAAAATTTGACCCCTTCAACCACTGATACATGAAAGCAGATTGGATGCCCTTGA 4026
D 3901 cagatcaaatcttgagaccccttcaaccagttacacgtgaagcagatttggatgtcccttga 3960
Q 4027 GAGGACACACATGAAAGAAATGATTGCTCAGCTACCTCTGAAACTTGAACTGAAAGTAT 4086
D 3961 gaggacacacatgaaagaatgatatgtctcagctacactcttgaacttgaaatcgtgaat 4020
Q 4087 GGAGATGGGATTAACCTCTCAGTGGGGGAACGGCAGCTCTTGTGACAGCTTAAGCCCT 4146
D 4021 gggagaatgggataaactctcagtggtgggaagcgagctctgtgataagtaagccct 4080
Q 4147 GCTCCGCACTGTAAAGATTGATTTAGATGAAGCCACAGCTGCCATGACACAGAGAC 4206
D 4081 gctccgcacagtaagatcttcatgatttagaagcccaagctgcacatgagacaaagagac 4140
Q 4207 AGACTTATGATTCAGAGACATCCGGAAGCAATTCACAGCTGATACATGAGTACCAT 4266
D 4141 agacttattgtatcacaagagacacacccgaagacatctgcagacgttacaatgctgagacat 4200
Q 4267 TGCCCATGCCCTGACACAGGTTTCAAGGCTCCGATAGATATGATGATGATGATGATGATGAT 4326
D 4201 tgcacatgcctgtgacaagagttctcaggtccgatagatgattatgtgtgtgtgtccagggaca 4260
Q 4327 GGTGTGAGATTGTGACACCCCATGCGTCTCTGTGCAACAGACAGTTCCGATTTCTATGC 4386
D 4261 ggtgtgtgagttgtgacaccccatcgttctctgttccaaagacagtttccgatcttctgtc 4320
Q 4387 CATGTTGCTGCTGAGAGCAAGGTGCTGTCAGAGGCTGACAGCTCCCTGTTGACGA 4446
D 4321 catgttctgtctgcaagaaagtcgtcgttcaagggcgaagctccctctgttgaaga 4380
Q 4447 AGTCTCTTTCTTTAGACATTTGCCATTCCTGCTGAGGGGGGGGCCCTTTCATGCGCTGC 4506
D 4381 agtctcttcttctttagagcatgtcmymkcbctgtgggggccccttcatcgtctc 4440
Q 4507 TCTTACGGAACCTTGCTTCTGATTTTATCTTTCGACAGACAGTTCCGATTTGGCTT 4566
D 4441 tcttaccgaaacctgtccttctcgaatttcttcttccagcagagtcgagattgtgctt 4500
Q 4567 GTGTGTTTCACTTTTAGGAGAGTGCATATTTGATTTGATTTATTTATTCATTTCAATGT 4626
D 4501 gtgtgttcaactttaaaggagatcaatcttgcataatgattatcaatcaatcaatcaat 4560
Q 4627 AAACAAATTTAGTTTGTCTTAAATGCACTCTAAAGGTTCAAGGAAACGTTATTTAT 4686
D 4561 aaacaaatttagtttcttcttcaatgtcaccttcaaaaggttcaagggaaacgttattat 4620
```


QY	4040	AAAAAATGTTTGGCTACGACTCCTCGAAACTTGAATCTGAAGTATATGAGAAATGGGAT	4039
Db	389	aaagaatgtaatctcaagctactccttgaacttgaatctgaagtatgtagaaatgtaggatt	448
QY	4100	AACATTCATATGGGGGAAGGAGCGCTTGTGCATATAGCTAGAGCCCTCTCGGACACTG	4159
Db	449	aactctcaatgtaggggaaggcagactcttgtagatagctagagccctctccgcaactgt	508
QY	4160	AAGATTCGTGATTTTAAATGAAAGCCACAGCTGCATGAGACACAGACACTTATGATT	4219
Db	509	aagatctgatttttagatgaagccacagctgcacatgagacagagacagactatttgaatt	568
QY	4220	CAGAAGACCATCCGAGAAGACATTTCAGACTGTACATGCTGATACCATTCGCCATGCTG	4279
Db	569	caagaagaccatccgaaagaacatttgaagactgtacacatgtacacattgccaatgcgctg	628
QY	4280	CACACGGTTCTAGGCTCCGATTAGATTAATGAGTGGGCCACAGGACAGTGGTAGATT	4339
Db	629	cacacggtctcagagctccgataagattatgtagctgagcccaaggacaggtgtagagctt	688
QY	4340	GACACCCCATACGGGCTCTTCGTGTCACAGCAGATTCGCCATTTTAAGCCATGTTTGCTCT	4399
Db	689	gaaccccatcgtctcctctcgtcccaagcagttccgattctatgcatgtttgctgct	748
QY	4400	GCAGAGAACAAGTCGCTGTCAAGGGCTACTCTCCCTGTTGACGAAGTCTTTTCTT	4459
Db	749	gcagaagacaaggtcgtgtcgaaggcttaccctccgttgaagaaagtcctttcttct	808
QY	4460	TAGAGCATTCGCAATTCCTCCGCGGGGGGGGCCCTCATGCGGCGTCCTACGGAAC	4519
Db	809	tagagcatctccatccctcgccctgggggggcccctccatcgctccctccacgcgaacc	867
QY	4520	TTCGCTTTCGATTTTATATCTTTCGCACAGCAGTTCGAGTTCGCTTGTTTCATT	4579
Db	868	ttgccttccgattttatctcttcgcacagcagttccgattgcttgtagtttcaactt	927
QY	4580	TTTAGGAGAGATCATTTTGTATATTGTATTTATTCATATTCATGCTAAACAAATTTAG	4639
Db	928	tttagaggagatcatatttgaattatgtattatcattatccattcatgtaaacaaaatttag	987
QY	4640	TTTTTGTCTTAAATGAGACTCTAAAAGGTTTCAGGGGAACGGTATTATTAATGTATACAG	4699
Db	988	tttttgccttaattgcaactctaaagggtccggggaaccggtattatataattgtatccag	1047
QY	4700	GCCATTAATGAAGCTTTATACGTGTAGCTATATCATATATATATTCTGTACATAACCTAT	4759
Db	1048	gccataatgaagccttatatacgtgtagctatatcatataataatcttgaactagcctat	1107
QY	4760	ATTTACAGTGAATAATGTAAGCTTTTATTTTATTTAAATTAAGACAGTGTCTAAAAAAA	4819
Db	1108	atttaccgtgaataatgtaagctgtttattcttattataataaagaacgtgtgtaataaca	1167

RESULT	8
AAV31498	
ID AAV31498 standard; cDNA; 5011 BP.	

DT	14-AUG-1998	(first entry)
XX		
DE	Human MRP variant ltpgpa (lei/pgpa) encoding cdna.	
XX		
KW	Multidrug resistance-associated protein: MRP; tumour; human; variant;	
KW	multidrug resistance; MDR; lei/smania P-glycoprotein; ltpgpa;	
KW	lei/pgpa; ds.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	196..4791
FT		/tag= a

FT	/product= "human MRP variant"
FT	2249
FT	variation
FT	/*tag= b
FT	/note= "a change from wild-type T to C"
FT	4039
FT	variation
FT	/*tag= c
FT	/note= "a change from wild-type C to G"
FT	4040
FT	variation
FT	/tag= d
FT	/note= "a change from wild-type G to C"
PV	US5766880-A.
XX	
PD	16-JUN-1998.
XX	
PF	05-JUN-1995; 95US-0463092.
XX	
PR	05-JUN-1995; 95US-0463092.
PR	27-OCT-1992; 92US-0966923.
PR	08-MAR-1993; 93US-0029340.
PR	26-OCT-1993; 93US-0141893.
PR	20-MAR-1995; 95US-0407207.
XX	
PA	(TOOH) UNIV QUEENS KINGSTON.
XX	
PI	Cole SP, Deeley RG;
XX	
DR	WPI; 1998-361687/31.
DR	P-PSDB; AAM57486.
XX	
PT	DNA encoding protein associated with multi-drug resistance - useful
PT	for as probe for identifying multi-drug resistant tumour cells
XX	
PS	Claim 12; Columns 67-78; 82pp; English.
CC	This cDNA encodes a variant of the human multidrug resistance-associated
CC	protein (MRP). This natural variant is a leismania P-glycoprotein related
CC	molecule ltpga (lei/pgpa). The human and murine MRP nucleic acid
CC	molecules can be used as probes for identifying multidrug resistant
CC	tumour cells by hybridisation to mRNA from tumour cells. The antisense
CC	nucleic acid can be used to reverse multidrug resistance (MDR). A
CC	recombinant expression vector containing the MRP nucleic acid molecules
CC	operatively linked to at least one regulatory sequence can be used to
CC	transform a host cell to produce a recombinant MDR-associated protein.
XX	
SQ	Sequence 5011 BP; 1064 A; 1498 C; 1394 G; 1055 T; 0 other;
<hr/>	
Query Match	10.5%; Score 511.2; DB 19; Length 5011;
Best Local Similarity	52.5%; Pred. No. 2e-121;
Matches 1367; Conservative	0; Mismatches 1178; Indels 59; Gaps 9;
QY	1811 CACATCCACTGGGCGCCTTACAGAGACACTGCACGATCATGTGAGATC 1870
DB	2142 cacatcaccttgaggcca--ggagcgacctccacactgaaatgcacccttcacac 2199
QY	1871 CAGAAGGTAAACTGTGGGAATCTGCGGCACTGTGGCAAGTGAAAACTCTCATTT 1930
DB	2200 ccggaagtgtcgttggtyggccgttgttgggccagtggyggtctgsgaaagtcgccctgc 2259
QY	1931 TCAGCCATTTTAAGGCCAGATGACGGCTTCTAGAGGGCGACGATTCANATCAGTGAACCTTC 1990
DB	2260 tcagccctcttgcctgtaagtggacaagaatggaggggcaagctgtcatcaaggtccgtg 2319
QY	1991 GCTTAGTNGGCCACGACGGCTGAGTCCGATGCTGACCTGAGAGACAACATCCGTGTT 2050
DB	2320 gcttatgtgcccacgacgaagcctggaatcgaatgatccctccgaaaacaacccctttt 2379
QY	2051 GGGAAGGATATGATGAGAAGAAAGATACAACCTGTGCTGCAACAGCTGCTGAGGCT 2110
DB	2380 ggatgtcagctgtggaagaaacatatatacaaggtcgtgatacagcgctgtgccccctcca 2439
QY	2111 GACCTGGCCATTTCTTCCAGCAGCACCTGCAGGAGANTGGAGAGCGAGGCCAACCTG 2170

Db 2440 gaacctggaacccctgcccagatcgagacagatctgaggaagggctgaacctg 2499
QY 2171 AGCGGTGGCAGCCGACAGAGATCAGCTTGGCCGGGCTTGTATAGACAGAGACATC 2230
Db 2500 tctggggagcaagaagcagcgctgagcctggccggcgctgactccaagcctgaact 2559
QY 2231 TACATCTCGACGACCCCTCAGTGCCTTACATGCCATGTGGGCAACACATCTTCAT 2290
Db 2560 tactcttcgaatgatacccttcacgctgagatgcaccatgtgggaaacacatcttga 2619
QY 2291 AGTGTAT-----CCGAAACATCTCAAGTCAAGACAGTTCGTTGTTACCCACAG 2344
Db 2620 aatctgactgcccccaaggagatgctgaagaacaagaacgagatctggtacacagac 2679
QY 2345 TTACAGTACCTGTTGATGATGAACTGATCTTCATGAAAGAGGCTGATTAACGGA 2404
Db 2680 atgagctactctgcgcagctgagactcaatcgtcaatgagctggcggaagatctcag 2739
QY 2405 AGAGGACCCCATGAGGAATGATTAATTAATGTGACTATGCTACCATTTTAAATAC 2464
Db 2740 atggctctcactacagagctgctgctcgagacgagccttcgagcttcctgaccc 2799
QY 2465 CTGTTGGTGGAGAGACACCCGCACT---GAGATCAATTCAAAAAGAAACAGTGT 2521
Db 2800 tatgccagcaagagcagagagagatgcaagagagaagcgggtcaacggcgtcagcgt 2859
QY 2522 TCACAGAGAAGATGACAGACAGAG-----GTCTTAAACAGATCAATTAAGAGG 2573
Db 2860 ccaggaagagagaaagcaaatgagaaatgcaatgcatgctgagcagacagcagggag 2919
QY 2574 AAAAAGCAGTAAACGACAGAGAGGAGCTTGTGACACTG----- 2615
Db 2920 caactgcagagacagctcagctcctccctcctatagtgagacatcaagcagccac 2979
QY 2616 AAGGAAAGGAGGAGGTTTCACTGCGCTGTGATGATATAGTGTACATCCAGGCTGTG 2675
Db 2980 aacgcagcagcagacagctgcaaaaagctgagcacaagaaggagagacctggaagctg 3039
QY 2676 GGGG-----CCCTTGGCATTCCTGCTTATTAAGCCCTTTTCATCTGAATGAGC 2728
Db 3040 gagctgcaagagcgacagcagcagcagctcagcttcctcgctgactggactcaatgag 3099
QY 2729 AGCACCCTTCACAGCCTGCTGTGTTGATCTGATCAAGCAAGAGAGGGAACACC 2788
Db 3100 gccactgcagactcttcactcctcctcagcatctcctcttcactgtaaacatgctgc 3159
QY 2789 ACTGTGACT-----CGAGGAGACGAGACCTCGGTGAGTGAAGATGAAGACATCT 2842
Db 3160 ggcctgctccaactatgctcagcctcagcctcagactgagcccccatcgtcaacggagc 3219
QY 2843 CATATGAGTACTATGCCAGCATCTACCCCTTCATGCATGCATGCTGATCCTGAAA 2902
Db 3220 caggagcaacagaaatgcctgagcgtlctatagagccctgggcatctcaacaagggtc 3279
QY 2903 GCCATTGAGG-----AGTTGTCTTGTCTCAAGGACGCTGACGCTTCTCCGGCTG 2956
Db 3280 gccctglttgctactcactcagtcgctgtccatcgggggatcttgctcccgctgctg 3339
QY 2957 CATATGAGCTTTTCCAGAGATGCTTCGAGAGCCCTATGAGTTTTCAGACAGCCCC 3016
Db 3340 cagctgacccctgcagacagatccctgcgtccacccatgagcttcttgcagcccccc 3399
QY 3017 ACAGGAGGATTTCAACAGAGTTTTCMAAGACATGATGATGATGATGCTGGCTGGC 3076
Db 3400 agtgggaacccgtggaacccgtctccaagaagctggaacaagcagactccatgctccg 3459
QY 3077 TTCCAGGCCAGATGTCATCCAGACGTTATCCTGCTTCTCTCTGTGGGAATGATC 3136
Db 3460 gaggtcatcaagatgctacatgagctccctgtccaagctcatgtgctgcatcgttcc 3519
QY 3137 GCAGAGAGCTTCGCGGTGTTCTTGTGGCAGTGGGGCCCTTGATCAGTCCTTTCAATC 3196

Db 3520 ctgctgcccagcccacatgcgcgcacatcatctccgcctctgctcactatctcttc 3579
QY 3197 CTGCACATTTGCTCCAGGGGCTCTGATTTCCGGAGCTGAAGCTGTGGACATATCACGAG 3256
Db 3580 gtccagaggtcttaacgtgtcttccctcccgagctgaagcctctcgatccgacccg 3639
QY 3257 TCACCTTCCTCCCATACATCAGCTCCAGCATPACAGGGCCTTGGCACCATCCAGCCATC 3316
Db 3640 tcccgcctactatcccatctcaacagagaccttgcgtggggtcagactcagactcc 3699
QY 3317 AATAAGGGCAGAGATTTCTCCAGACATACAGAGAGCTGCTGATGACAAACACTCTCT 3376
Db 3700 gagagagcagagcgcttcatccacagagtgactcctgaagcagcagaaacacagaagcc 3759
QY 3377 TTTTTTTGTTTAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3436
Db 3760 tattccccagatcgtgtgcaacagctggtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 3819
QY 3437 GCCCTATCACACACAGGGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 3496
Db 3820 tgcactgctctgttctgctgcctgttgcgtgagatcctcagacacagcctcagctgagc 3879
QY 3497 TATGGGGTCTGCGCCATCTTATGCTGTGATGATGATGATGATGATGATGATGATGATG 3556
Db 3880 ttggtggcctctcagtgcttactcactcagtgacgttcacacagcttgaactgctggt 3939
QY 3557 AGACTGCATCTGACACAGAGCTCGATTCATCTCGGTGAGAGATCAATCTACATTT 3616
Db 3940 cgaatgtcatctgaatgaaacacacatcgtggcgtgagagcgtcaagaagatctca 3999
QY 3617 AAGACTGTGCTGTTGAGAGACCTGCCAGATTTAAGAACAGAGGCTCCCTCCGATCGG 3676
Db 4000 gagactgagaa---ggaagcgccctgcaaatccaggaagaagctccgccagcagctgg 4056
QY 3677 CCCGAGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3736
Db 4057 ccccaagtgagcgagtgatctccgagactacgtcgcgtcagcagagagactggac 4116
QY 3737 CTGCTCTTAAAGAAATGATCTTCAAGATCAAACTTAAAGAAAGATTTGCAATTTGGGG 3796
Db 4117 ttgcttccagcgacatcaatgatacagatcagatgaggaagaaagctgcatcgtggg 4176
QY 3797 CGGACGAGTACGAGGAATCTCGCTGGGGATGGGCCCTTCCGCTGCTGTGAGATTAATCT 3856
Db 4177 cggagggagctgggaagctgctccctgacccctggccttalttcgatacagagctgc 4236
QY 3857 GGAGGCTGACATCAAGATTTGATGAGTGAATGATGATGATGATGATGATGATGATGATG 3916
Db 4237 gaagagagatcatcatcgtatgagatcaacatcgccaagatcgcgcctgcagactccgc 4296
QY 3917 AGCAAACTCTGTATCATTCCTCAAGAGCGGCTGCTGTTCACTGACATGTCAGATCAAT 3976
Db 4297 ttcaagatcacatcatcccccagagccctgttctgttcgggttccctcagatgac 4356
QY 3977 TTGAGCCCTTCACACATATACATGAAACACAGATTTGGATGGCCTTGAAGAGACAC 4036
Db 4337 ctggaaccatccacgaactacgatacgaatgaagaagctcggagatccctcggagctgc 4416
QY 4037 ATGAAGATGATTTGCTGACCTGACCTGATGATGATGATGATGATGATGATGATGATG 4096
Db 4417 ctgaaagactcgtgctgacgccttcctcgtacaaagctagaacatgaaatgacgaaggg 4476
QY 4097 GATAACTCTCATGTTGGGGGAAGCGAGCTCTTGTGATAGCTAAGCCCTGCTCCGCCAC 4156
Db 4477 gagaacctcagtgctggcgagcgcagctgtgtgctcagccggccctgtcagagag 4536
QY 4157 TGTAAAGATTTGATTTTGAAGAACACACACTGCGATGGAACACAGAGACGACTTATG 4216
Db 4537 acgaagatcctgtgtgagatgagcagcagcagcagcagcagcagcagcagcagcagc 4596
QY 4217 ATTCAAGAGACCTCCGAGAGATTTTCAAGATGATGATGATGATGATGATGATGATGATG 4276
Db 4597 atccagttcacacatcccgagacagatcggagactgacagctccatccatccatccac 4656

OY	4277	CTGACACGGTTTCAGCCCTCGATPAGATTATGTGCTTGACCACGAGACAGTGTGGAG	4336
Db	4657	ctcaaccatcatgatgactacaagaagggtacgttccttgcagacaaggagaalccagag	4716
OY	4337	TTTGACACCCCATCAGGTCTCTTCTG	4360
Db	4717	tacggcgcccatcgcagacctcgt	4740
RESULT 9			
ID	AXI19818	standard; cDNA; 5011 BP.	
XX	AXI19818;		
XX	AAAI9818;		
DT	10-JUN-1999	(first entry)	
DE	Human multidrug resistance-associated protein natural variant cDNA.		
XX	Human; multidrug resistance-associated protein; MRP; cytotoxic drug;		
KW	cancer; chemotherapy; ds.		
XX			
OS	Homo sapiens.		
Key	Location/Qualifiers		
FH	CDS	196..4791	
FT	/tag= a		
PN	US5891724-A.		
XD	06-APR-1999.		
XX			
PF	05-JUN-1995;	95US-0460907.	
XX			
PR	05-JUN-1995;	95US-0460907.	
PR	27-OCT-1992;	92US-0966923.	
PR	08-MAR-1993;	93US-0029340.	
PR	26-OCT-1993;	93US-0141893.	
PR	20-MAR-1995;	95US-0407207.	
XX			
PA	(TOOH) UNIV QUEENS KINGSTON.		
PI	Cole SPC, Deeley RG;		
DR	WPI, 1999-253868/21.		
DR	P-PsDB; AAW93894.		
PT	Protecting mammalian cells against cytotoxic drugs		
XX			
PS	Claim 4; Column 67-78; 82pp; English.		
XX			
CC	The present sequence encodes a human multidrug resistance-associated		
CC	protein (MRP). The present invention also describes a method for		
CC	protecting a mammalian cell against the cytotoxicity of anthracyclines,		
CC	epidoxophyllotoxins and Vinca alkaloids (A) by introducing into it a		
CC	nucleic acid (I) that hybridizes under stringent conditions to a nucleic		
CC	acid (II) that encodes an MRP protein (III). Introduction of (I)		
CC	protects cells against cytotoxic effects of (A), particularly to protect		
CC	normal cells against (A) being used for treatment of cancers..Cells		
CC	transformed with (I) can be used to screen for agents that affect		
CC	multidrug resistance or are directly toxic to multidrug resistant cells,		
CC	i.e. potential therapeutics for multidrug-resistant cancers. Confering		
CC	resistance to normal cells should allow an increase in the dose of (A)		
CC	that can be administered safely.		
XQ	Sequence 5011 BP; 1064 A; 1498 C; 1394 G; 1055 T; 0 other;		

Query Match	10.5%;	Score 511.2;	DB 20;	Length 5011;
Best Local Similarity	52.5%;	Pred. No. 2e-121;		
Matches 1367;	Conservative	0;	Mismatches 1178;	Indels 59; Gaps 9;

[illegible]

QY 2843 CATATGCAGTACTATATGCCAGCATCTACGCCCTCTCCATGGCAGTCATGCTGATCTGA 2902

Db 2561 taaccttcgatgatcccttcacgaagtgatgcccacgtggtggaaacaacatctttgaa 2620
QY 2291 AGTGCAT-----CCGAAACATCTCAAGTCCAGACAGTTCTGTTTGTACCCACG 2344
Db 2621 aatgtagtgcgcccagaagtgatgctgaagaacaacgcgcgaatctgtgtccgcacagc 2680
QY 2345 TTACAGTACCTGTTGACGTGATGAAGATCTTCATCAAGAGGCGGTATTACGGA 2404
Db 2681 atgaagctactgcgcgaagtgacgtcaltcatcgtcactgagtcgcgaagatctctgag 2740
QY 2405 AGAGACCCCATGAGAGACATGATGAATTTAAATGAGTATGCTACATTTTAAATAC 2464
Db 2741 atgggtccttaccaggaagtgctgtcgaagcgcgttcgtgagtctctgtgtaac 2800
QY 2465 CTGTTGCTGGAGACACCCGCACTT---GAGATCAATTCAAAAAGAAACCACTGAT 2521
Db 2801 tatgcacagcacagacaggaatgcagagtgaggaacggtgtccaggtcgctcagcgt 2860
QY 2522 TCACAGAAAGTACACAAAGCAAG-----GTCTTAAACAGGATCAATTAAGAGG 2573
Db 2861 ccaggaagaagaagcaagaatggaatgacatgcatggtgacgagtcaggtgacggtgag 2920
QY 2574 AAAAGCATTAAGCCAGAGAGGCAAGCTGTGCAGCTGG----- 2615
Db 2921 caactgcagagacagctcagcagctcctcctcctactatagtggtgaaactcagcagcaac 2980
QY 2616 AAGAGAAAGGACAGGTTGATGCCCCGTGACATGATATGCTGTCAATCCAGCTGCTG 2675
Db 2981 aacacacacgcagaaactgcgaagaactgcagcaagaagagagacccgtgagctgtagt 3040
QY 2676 GGGGG-----CCCTTGCAATTCCTGTTATTATGGCCCTTTTCAATGCTGATGAGGC 2728
Db 3041 gaggctgacaagcgccagacagagcgacgtcctccgtgctacgggtactacaagaa 3100
QY 2729 AGCACCGCTTCAGACACTGTGTGTTGATTTACTGTGATCAACCAAGAGAGCGGAACACC 2788
Db 3101 ggcatacgcagctctcatctcctcctcagcatctcctcttcaatgtgatacaatgctgc 3160
QY 2789 ACTGTGACT-----CGAGGGAACGAGACCTCGGTGATGACAGCTGTAAGACAACTCT 2842
Db 3161 ggcgtgcgtctccaactatgtgtcagctcctctgactgatacgaccacatgcacaggtact 3220
QY 2843 CATATGACATGATATGCCAGCATCTACGCCCTCTCCATGGACGTGCTGATCTGATA 2902
Db 3221 cagggacacacgaagaatccggtcgtgactatgagccctggacatccacaaggtatc 3280
QY 2903 GCCATTCGAG-----AGTTGTCTTTGTCAAGGCGACGCTCGAGCTTCCTCCGGCTG 2956
Db 3281 ggcgtgttctgtactcactcgtgcgtgtcaltcgggggtactgtgtccgcgtgtctg 3340
QY 2957 CATGACGAGCTTTTCCGAAGGATCCTTCGAAGCCCTATCAAGTTTATTACAGACCCCC 3016
Db 3341 caegtggacctgtgcacagcatcctcgtgcacacatgaagctctcttgaagggagacccc 3400
QY 3017 ACAGGAGAGATTCTGAACAGGTTTCCAAAGACATGATGATGATGACGTGGCGCTGCC 3076
Db 3401 agtggaaacctgtgaaacctcttccaaagagctggacaacgttggatccatgatcccg 3460
QY 3077 TTTCAGGCCGAGATGTTCAACAGACGTTATCTGCTGTTCTTCTGTTGGAATGATC 3136
Db 3461 ggggtcatcaagatgttcatgtggctccctgttcaagcatcatgtgtgctcgtatcc 3520
QY 3137 GCAGAGAGTTCCTCCGAGTCTCTGTGGCAGTGGGGCCCTGTGATCCTCTTTTCAGTC 3196
Db 3521 ctgtgtgcacagccatcgcgcacatcatcaccgcccttggctcctactactcttc 3580
QY 3197 CTGACATGTTCTCCAGGTTCTGATTTGGGAGCTGAACGCTGTGACATATTCACGAG 3256
Db 3581 gtccagaggttctaagtgtcttccctccggcagctgaaagcgctcgtgacgtcagcgcg 3640
QY 3257 TCACCTTCTCTCCACATCAAGCTCCAGCATACAGGCGCTTGCCACCATTCACGCGCTAC 3316
Db 3641 tcccggtctattcccatccaacgaagacctgtgcgtgggtgcaagcgtcatctgaagcttc 3700
QY 3317 AATTAAGGCGAGAGTTTGTGCACAGATACAGAGACTCTGTGATGACAAAGCTCCT 3376
Db 3701 gagagcagaagacgtctcatccaacagagtgacctaaggttgaacgaagacagaagggcc 3760
QY 3377 TTTTTTTTTTACGTGTCGATGGCGTGGCTGGCTGTGCGGCTGGACCTCATACATAC 3436
Db 3761 taattcccacgaatcgtgtgcacaagatggtgcgcgtgcgtgcgtgagtgtgtggcaac 3820
QY 3437 GCCCTCATCCACACACGCGGCTGATGATGTTCTTATGTCACGGGAGATTCCCCAGCC 3496
Db 3821 tgcattcgtctgttgcctcctcgttctgttgatctccagccaagctcagttgtgtgc 3880
QY 3497 TATCGCGTCTGCCATCTTATGCTGTCCAGTTTAACGGGCTGTGCCAGTTTACGGTC 3556
Db 3881 ttgtgtgcctctcagttgcttactactatcagttacacacgtactgtgactgtgtgt 3940
QY 3557 AGACGTGATCTGAGACAGAGCTGATTCACCTCGGTGGAGAGATCAATCATCTCATTT 3616
Db 3941 cggatgtcatctgaaatggaacccaacatcgtgcgtggaagagctcagaagatatta 4000
QY 3617 AAGACTGTCTCTTGGAGACACCTGCGACAAATTAAGAACAGGCTCCCTCCCTGACTGG 3676
Db 4001 ggaactgagaa---ggaggcgcccttgcaaatccagtgagacagctccgcacagcgtg 4057
QY 3677 CCCAGAGAGGAGAGGTGACTTTTGAAAGCAGAGATAGTACCGAANAACCTCCT 3736
Db 4058 ccccaagtgtgcagatggaatctcgaactactgtcgtcactcgaagagaaactgcgac 4117
QY 3737 CTGCTCTTAAAGAAAGTATCTTTCACAGTCAAACTTAAGAGAAATGCGATTGGGG 3796
Db 4118 ttcgttctcaaggacatcaatgtcaagatcaatbgyggaaggtcggactcgtggtg 4177
QY 3797 CGAGACGATCAGGAAGTCTTCGCTGGGAGTGCCTCTTCCTGCTGTGAGATTATCT 3856
Db 4178 cggagagagcttgggaagtcgtccctgacccctggttatttcogtaacagatctgc 4237
QY 3857 GGAAGCTGATCAAGTTTATGAGTGAATTCAGATTCAGTATGCTTCCGACCTCCGA 3916
Db 4238 gaagagagatcatcatcgtatgcatcaacatcgcgaagatcgcgtcgaagcctccgc 4297
QY 3917 AGCAAACTCTATCATCTTCCTCAAGAGCGGTGCTGTAGTGGCACTGTCAATCAAT 3976
Db 4298 ttcaagatcacatcatccccaagacccgtgttgttgcgttccctccogaagaa 4357
QY 3977 TTGACCCCTTCAACCATGATCACTGAAAGCACAGATTTGGGATGCCGAGACAC 4036
Db 4358 ctggaccatctcagccagatcgtgataagagtctgagcgtccctggagctggtccac 4417
QY 4037 ATGAAGAATGTATTGCTACGTACTCTGAACCTTGAACTTGAACTGAAATGGG 4096
Db 4418 ctgaagagacttgcgtcagcccttctgacaagctagacatgaatgtgcagaagcggtg 4477
QY 4097 GATTACTTCTAGTGGGAGGAGCGCAGCTCTTGATGATGATGAGAGCTGCTCCGAC 4156
Db 4478 ggaagacctgaigtgcggaagcgcaagcttgtgtgctatgcccggccctgtgaggaag 4537
QY 4157 TGTAAATTTGATTTTATGATGAAGCCACAGCTGCCATGACACAGACATTTATTG 4216
Db 4538 acgaatcctctgtgtgtgtagtgcgcagcgccgtgacctgtgaaacgagccctc 4597
QY 4217 ATTGAAGAACATCCGAGAAAGCATTTTGCAGACTGTACATGCTGACATTCCTCC 4276
Db 4598 atccagtcacacatccggacacagatgcgaagactgcgacgtccctcaaccatcgcac 4657
QY 4277 CTGACAGGTTCTAGGCTCCGATAGGATTAATGCTGCTGCCACGAGACAGTGGTGA 4336
Db 4658 ctcaacacatcatgtgactcaacaaggtgtagtcttcttgacaagaaggaatccaggg 4717
QY 4337 TTTGACACCCATTCGCTCTTG 4360
Db 4718 tagggcccatcgtgacctctg 4741

```
RESULT 12
AA290193
ID AA290193 standard; cDNA; 5011 BP.
XX
AC AA290193;
XX
DT 19-MAY-2000 (first entry)
XX
DE Multidrug resistance protein (MRP) natural variant nucleotide sequence.
XX
KM Multidrug resistance protein; MRP; human; anthracycline; Vinca alkaloid;
KW epipodophyllotoxin; cancer; Leukaemia; variant; ss.
XX
OS Homo sapiens.
XX
PN US6025473-A.
XX
PD 15-FEB-2000.
XX
PF 05-JUN-1995; 95US-0461384.
XX
PR 27-OCT-1992; 92US-0966923.
PR 08-MAR-1993; 93US-0029340.
PR 26-OCT-1993; 93US-0141893.
PR 20-MAR-1995; 95US-0407207.
XX
PA (TOOH ) UNIV QUEBENS KINGSTON.
XX
PI Cole SPC, Deeley RG;
XX
WPI: 2000-181838/16.
DR P-PSDB: AAY78873.
XX
PT Isolated protein conferring multidrug resistance, to at least two drugs
PT selected from anthracyclines, epipodophyllotoxins and Vinca alkaloids,
PT on a drug sensitive mammalian cell -
XX
PS Disclosure; Column 69-80; 78pp; English.
XX
CC This sequence represents a human multidrug resistance protein (MRP)
CC natural variant nucleotide sequence. The human MRP encoded by this
CC sequence and that encoded by wild-type MRP (see AA290192) confers
CC multidrug resistance, including resistance to at least two drugs selected
CC from anthracyclines, epipodophyllotoxins and Vinca alkaloids, on a drug
CC sensitive mammalian cell, when the protein is expressed in the cell. The
CC multidrug resistance is not substantially reversed by chemosensitizers
CC which reverse P-glycoprotein-mediated multidrug resistance. The MRP
CC protein sequence can be used to generate antibodies against MRP. The MRP
CC protein and nucleotide sequences can be used in compositions which are
CC used to treat patients with tumours displaying multidrug resistance. The
CC compositions and methods of the invention can be used particularly to
CC treat breast cancer, leukemias, fibrosarcomas, cervical cancer,
CC gliomas, thymomas, neuroblastomas, and lung cancer. Antibodies directed
CC against MRP can be used to inhibit the multidrug resistance of a
CC multidrug resistant cell.
XX
SQ Sequence 5011 BP; 1064 A; 1498 C; 1394 G; 1055 T; 0 other;
```

```
Query Match 10.58; Score 511.2; DB 21; Length 5011;
Best Local Similarity 52.58; Pred. No. 2e-121;
Matches 1367; Conservative 0; Mismatches 1178; Indels 59; Gaps 9;
```

```
1811 CACATCCACCTGGGCGACCTGCTACAGAGACATGCACAGATGATGGAGATC 1870
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2142 caacttaaccgggcca -ggagcgacctccacacggaatgcatcaacttcac 2199

QY 1871 CAAAGAGGTAACGCTGTGATCTGGCGCAGTGTGGAGAGTGAACCTCTCTCAT 1930
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2200 ccggaagtgcttggtggtggtggtggtggtggtggtggtggtggtggtggt 2259
```

```
QY 1931 TCAGCCATTTAGGCCAGATACGCTTCTAGAGGGCAGCATTTGCAATCAGTGAACCTTC 1990
||||| | | | | | | | | | | | | | | | | | | | | | | |
Db 2260 tcagccctctggtcgtgagatgacaaagtggagggcgacgtggatcaaggtccgtg 2319

QY 1991 GCTTATGTGGCCAGCAGAGCGCTGATCCTCAATGCTACTCTGAGAGCAACATCTGTTT 2050
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2320 gctatgtgccacagcagggcttgatcagaatgattctcccgagaaaaatccctt 2379

QY 2051 GGAAGAAGATATGATGAAGAAAGATACAACTCTGTCTGAACACCTGCTGAGGCTT 2110
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2380 gtagtgcagctggaggaaccatattacaggtccgtgatacagcctgacctccca 2439

QY 2111 GACCTGGCATTTCTCCAGCAGCAGCAGCAGCAGATGTGAGAGCAGAGCCACCTG 2170
||||| | | | | | | | | | | | | | | | | | | | | | | |
Db 2440 gacctggaatccctgcgcaggtgggtgatcgagacagatgtgcgaaaggtgaa 2499

QY 2171 AGGGTGGGCGAGCGCCAGAGATCAGCCTTCCGGGGCTGTATGAGACGAGGATC 2230
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2500 tcggggagacagaagcagcggtgagcctggccggcggtgactccaacgtgacat 2559

QY 2231 TACATCCGAGCAGACCCCTCAGTGCCTTAGATGCCATGTGGCAACACATCTTCAT 2290
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2560 tacctctcgatgataccctctcagcagtgatgccatgtgggaaaaacacatcttga 2619

QY 2291 AGTGTAT-----CCGAAACATCTCAAGTCCAAAGACAGTTCCTGTGTACCCACAG 2344
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2620 aatgtgaltggcccaaggagatgtcgaagaacaagacgagatcttgytacacagc 2679

QY 2345 TTACAGTACTGCTGTTGACTGTGATGATGATGATGATGATGATGATGATGATGATG 2404
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2680 atgagctacttgccgaggtgagatcatalcgtcagtgatgagtgagcaaatctcga 2739

QY 2405 AGAGGACCCCATGAGGAAGCATGATGATTAATGATGATGATGATGATGATGATGATG 2464
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2740 atggctctctacagaggtgctgctgagagcggcttcgagttctcgtac 2799

QY 2465 CTGTTCGCGAGAGACACCGCCAGTT---GAGATCAATTTAAAAAAGAAACCAAGTGT 2521
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2800 tatgccagacagagcagagatcagatcagagagagagagagagagagagagag 2859

QY 2522 TCACAGAGAGAGTACAGACAGACAGG-----GTCTAAACAGAGATCAATTAAGAGAG 2573
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2860 ccagggagaggaagcaagcaaatggaatgcatgctggtgagcaggaagtcagggag 2919

QY 2574 AAAAAAGTAAGCCAGAGAGAGAGAGGAGCTGTGTGACAGTGG----- 2615
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2920 caactgcagagacagctcagcagctcctccctcctcctcctcctcctcctcctc 2979

QY 2616 AAGAGAAAGGCGAGGCTGATGCTGCTGATGATGATGATGATGATGATGATGATG 2675
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2980 aacgcaacgcagagacgtcagaagaagctgagggcagaagagagagagagagag 3039

QY 2676 GGGGC-----CCCTTGCAATCTCTGCTGATGATGATGATGATGATGATGATG 2728
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3040 gaggctgcaagagcgagcagagagagagagagagagagagagagagagagagag 3099

QY 2729 AGCAGCGCTTCAGACACCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 2788
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3100 gcaatcgagactctcatctcctcctcctcctcctcctcctcctcctcctcctc 3159

QY 2789 ACTGTGACT-----CGAGGAGACGAGACCTCGGATGATGATGATGATGATGATG 2842
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3160 ggcgtgacttcaactatgtgctcagcctcctgagctgagctcccatcgtcaagggag 3219

QY 2843 CATATGCACTATGTCAGAGATCTACGCTTCATGAGAGATGATGATGATGATGATG 2902
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3220 cagagacacagaaagtcgcgtgagcgtcatgtagagccctgggcatltaacaaggatc 3279

QY 2903 GCCATTCGAGG-----AGTTGCTTTGTCTAAGGAGACGCTGCGAGCTTCTCCGCGTG 2956
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3280 ggcgtgttgctactcactgacgtgctgctcagtcgagggagagagagagagagag 3339

QY 2957 CATGACGAGCTTTTCCGAGAGATCTTCGAGGCCCTATGAGAGTTTTCGACAGACCC 3016
```

```

Db 3340 caagcggacgtctgacagcatcctcggtcacccatgagctctcttgagcgagacccc 3399
OY 3017 ACAGGAGGATCTCAACAGGTTTCCAAAGACATGAGTGAAGTACCTGGGCTGCCG 3076
Db 3400 agtggaaacctggtaaacgctctcccaagagctggacaacatgacatccatccg 3459
OY 3077 TTCAGGCGGAGATGATCTCAACAGGTTTACGTCGTGTCTTCTGTCGTGGGAATGATC 3136
Db 3460 gaggtcatcaagaatgttcatctggtccctgttcaacgltcatctggtcctgcatc 3519
OY 3137 GCAGGAGCTTCCCGTGGTTCCTTGTGGCAGTGGGGCCCTTGTACCTCTTTTCAGTC 3196
Db 3520 ctgcgycacagcccatcgccgcatcatcatccgccccttggtccatctactcttc 3579
OY 3197 CTGCACATTTGTCTCCAGGGTCTTGATTCGGGAGCTGAAGCGTTCGACAAATATCAGCAG 3256
Db 3580 gtccagaaagtcttaacgttgcctcccgagctgaaagcgctcgaagtcggtcagcgcg 3639
OY 3257 TCACCTTTCCTCCACATCAGTCAGCATACAGGGCTTGGCCACATCCAGCGCTAC 3316
Db 3640 tccccggtcatctccatctcaacagacatctgctggggtcagcgltcatccgaagc 3699
OY 3317 AATTAAGGCGAGAGTCTTCTGCACAGATACAGAGCTGCTGATGACAAACAGCTCCT 3376
Db 3700 gagggagcagagcgcttcatccacagagtgagcctgaagtgagcagagacaagaagc 3759
OY 3377 TTTTGTGTGTACGTGCGATGCGGTGGCTGCTGCTGCGCTGACCTCATACAGATC 3436
Db 3760 taataccacagatcgltgcacagtggtgcgtgcgtgagtggtgtgtgtgtgtgtgt 3819
OY 3437 GCCCTATACACACACACGGGGCTGATGATCCTTATACAGGGGAGATTCGCCAGCC 3496
Db 3820 tgcacgtctgtctgtcgcctgtgttcggtgcatctccagcaagcctcagtggtgc 3879
OY 3497 TATGCGGCTCTGCGCATCTTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3556
Db 3880 ttggtggcctctcagtggttcatcatctgacagtgatccacagtaactgagctggt 3939
OY 3557 AGACTGCGATCTGACAGACAGAGCTGATTCACCTCGGTGAGAGATCAATCAATCATT 3616
Db 3940 cggatgtcatctgaaatgaaacacacatcgtgcgtgtgagaggtcgaagagatcat 3999
OY 3617 AAGACTGCTGCTTGGAGGACCTGCCAGATTAAGAAACAGGCTCCCTGACTGCTG 3676
Db 4000 gagactgagaa---ggaggcgccctgcaaatccaggaacagcctccgcccagcgtcg 4056
OY 3677 CCCAGAGAGGAGAGGTGACCTTTGAGAACGACAGATGAGGTACCGAGAAAACCTCCT 3736
Db 4057 ccccaagtggtgcgagtggaatctcggaactatgctcgtcgtacccgagagagcctgc 4116
OY 3737 CTGCTCTTAAGAAAGTATCTTCAACGATCAAACTAAAGAGAGATTTGCTGGG 3796
Db 4117 ttctgtctcaagacaatcatgcaatgcaatcgaatggtgggaaaggtcggtcggtgg 4176
OY 3797 CGGACGAGATCAGGAGAGCTCGCTGGGGATGGGCCCTCTCCGTCGTGGAGTATCT 3856
Db 4177 cggacggagagctcggaagtcgctcccgaccccggtctatctccgaatcaagacgtctgc 4236
OY 3857 GGAGGCTGCATCAAGATTGATGAGTGAAGTCAAGTATTTGCTTCCGACTCCGA 3916
Db 4237 gaagagagatcatcatctgcatcaatcatcgcaagatcggtcgtcgaagcctccgcg 4296
OY 3917 AGCAAACTCTATATCTTCTCAAGACCGGCTGCTTCAAGTGCGCACTTCAGATCAAT 3976
Db 4297 ttcaagatcacatcatcccccagagaccctgttctgttttcgggttccctccgaatgaa 4356
OY 3977 TTGACACCCCTTCAACAGTACACTGAAGACAGATTTGGATGCTCGAGAGACAC 4036
Db 4357 ctggaacccatcaagcagtaactcggatgagaagctcgtgagctcgtgagccccc 4416
OY 4037 ATGAAAGAAATGATTTGCTTCAGCTACCTCTGAAGCTTGAATCTGAAGTGAATGGG 4096

```

```

Db 4417 ctgaagagcttcgtgtcagccctctcgaacaagctagacatgaaatgtgcagaagcg 4476
OY 4097 GATPACTTCTCATGTGGGGAGGAGGAGCTTGTGATGATGATAGAGCCCTGCTGCCAC 4156
Db 4477 gagaacctgaatgtgcggagagccagctctgtgtcagcccggtccctgtgaaag 4536
OY 4157 TGTAAATCTGATTTTATATAGACAGCAGCTGCCATGACACAGAGACAGACTTATTG 4216
Db 4537 acgaagatcctctgtgtgtgtgtagagcgaagcagcggtgagactggaacgagacctc 4596
OY 4217 AATCAAGAGACATCCGAGAGACATTTGCAAGCTGATCAGTCCATGACATTCATCCAT 4276
Db 4597 atccacatccatcccgagacacagttcgaggaatgaccggtctcctcaacgcgcccag 4656
OY 4277 CTGCACACGCTTCTAGGCTCCGATGAGATTTATGTCGTGCTGCTGCTGCTGCTGCTG 4336
Db 4657 ctcacacatcatatgacatcacaaagtgatcgtcttggacaagaagaaatccagag 4716
OY 4337 TTTGACACCCCATCGGCTCCTTG 4360
Db 4717 taagcgcctccatcggaactctgt 4740

RESULT 13
AA239556
ID AA239556 standard; cDNA; 5011 BP.
XX
AC AA239556;
XX
DT 28-FEB-2000 (first entry)
XX
DE Human MRP variant encoding cDNA.
XX
KW Chemosensitizer; multidrug resistance-associated protein; MRP; human;
KW therapeutic agent; P-glycoprotein-mediated multidrug resistance; lung;
KW cancer; variant; ds.
OS Homo sapiens.
OS Synthetic.
XX
FH Key location/Qualifiers
FT CDS 196..4791
FT FT /*tag= a
FT mutation /*product= "MRP variant"
FT 2249
FT /*tag= b
FT mutation /*note= "wild-type T is changed to C"
FT 4039
FT /*tag= c
FT mutation /*note= "wild-type C is changed to G"
FT 4040
FT /*tag= d
FT mutation /*note= "wild-type G is changed to C"

US6001563-A.
PN
PD 14-DEC-1999.
XX
PF 05-JUN-1995; 95US-0463179.
XX
XX 27-OCT-1992; 92US-0966923.
PR 08-MAR-1993; 93US-0029340.
PR 26-OCT-1993; 93US-0141893.
PR 20-MAR-1995; 95US-0407207.
XX
XX (TOOH ) UNIV QUEBENS KINGSTON.
XX
PI Cole SP, Deeley RG;
XX
XX WPT: 2000-061877/05.
XX
XX P-PSDB; AA155799.
XX
PT Identification of chemosensitizers useful for treating cancer, using

```


Db	2233	ggatgctacgctgtaggaacacattaccaggtccgtyatccaggtcttgcctccctccca	2232
Qy	2111	GACCTGGCACTTCTTCCACAGCAGCCTGAGGAGATTGAGAGCGAGAGCCAACTG	2170
Db	2293	gactcgtgaatccctgcgccagctggggatcgcgaagagatttggagagaagggcgtgaacctg	2352
Qy	2171	AGCGGTGGGACAGCCACAGAGGATCAGCTTTGCCCGGGCCTTTGTAATAGTCAGAGAGCATC	2230
Db	2353	tcctggggagacgaagcagcgctgagacgtctggccgggcccgtgtactccaaagcttgcatt	2412
Qy	2231	TACATCTCGGAGCAGCCCCCTCAGTGGCTTAAGATGACCAATGTGGGACACACTTTCAT	2290
Db	2413	tacctcttgatgatccctccctcctaagcagtgatgtgccatgttgysaaacacattctttaa	2472
Qy	2291	AGTGTAT-----CCGAAACATCTCAAGTCCAAACACAGTCTGTGTATACCCACAG	2344
Db	2473	aatgtgattggcccccaaggatgtcgtgaagaacaaagccgcatcttgttaacgaacgc	2532
Qy	2345	TTACAGTACTCGTGGTACTGTGATGAAGTATCTCATGAAGAGCGCTGATTAACGSA	2404
Db	2533	atgagctacttgcgcgcaggtgcagtcacatcgtatcgtatgagtgysgagaagatccttga	2592
Qy	2405	AGAGGACCCACTGAGCAACTGATGATTTAAATGGTGATATGATCACCACTTTTAAATAC	2464
Db	2593	atgggtccctaccagagcgctgcttgctcgagaagcgccctgcgtgagttccgtgcacc	2652
Qy	2465	CTGTGCTGGAGAGACACGCCCACTT---GAGATCAATTTCAAAAAGAAACAGTGGT	2521
Db	2653	tatgcacagcacagcgcaagcgcaagatgtcgaagagaacgggtcaacggcgctcaagcgct	2712
Qy	2522	TCACAGAGAAGTCAACAGCAAG-----GTCTTAAACAGAGATCATTAAGAAGG	2573
Db	2713	ccaggaagaaagaaagcaaaatgagaaatgcatgctgtgtacgcaagctgcagysaag	2772
Qy	2574	AAAAAGCAGTAAACCCAGAGAGAAAGGAGAGCTGTGCAGTGG-----	2615
Db	2773	caactcgaagaaacagcgtccagcagctcctcctcctaatagtgysgacatcaagcagcaacc	2832
Qy	2616	AAGAGAAAGGGCAGGAGTTTCAGTGCCTGTGATGATATATGATGTCTACATCCAGCTGTG	2675
Db	2833	aacagcaacgcgaagacgtcagaaagcttgagcccaagaagagagaccttgysagcttgatg	2892
Qy	2676	GGGG-----CCCTTGGCATCTTCGTGTTATATGAGCCCTTTTATGCTGAATAGGC	2788
Db	2893	gagcgttacaaagygcgacgaagagcaggtatgaacttccgtgtactgysgactataagaa	2952
Qy	2729	AGCACCGGCTTACACACCTCGTGTTGAGTACTGATGAACAGCAAGAAAGCGGAAACAC	2788
Db	2953	gcatcggaactcttcaactcctcctcccaagatctccttcaigtgtaaacatgttcc	3012
Qy	2789	ACTGTGACT-----CGAGGGAACGAGACCTCGTGAGTGAAGAGATGAAGACAACTCT	2842
Db	3013	gcgcgtgcttccaaactatttgcctcagcctcttgacgtatgagcccatcgttaacgysgact	3072
Qy	2843	CATTTGCACTACTATGCGACAGCTTAAGCCCTCTCCATGGCAGTACGTGATCTGTAA	2902
Db	3073	caggagcaacgaagaaagtcgcgctgagcgtctatagagcccttggacttcaacaagysatc	3132
Qy	2993	GGCATTTGAGG-----AGTTGTCTTTGTCAAGGAGACCTCGAGCTTCTCCCGGCTG	2956
Db	3133	gcgcgtgttgcctactccatcgcgcgtgtcatcgtgggggagatcttgcttccgcgctgtctg	3192
Qy	2957	CATACAGAGCTTTCCAGAGATCTTTGCAAGGCCCTATGAATTTTGTGACAGCAGCCCC	3016
Db	3193	caagtgaactgtcgtcagcagcatcctcgtggtcaaccatgaagcttcttggagysgaacccc	3252
Qy	3017	ACAGGAGAGATTCTCAACAGGTTTCCAAAGACATGATGAAGTTGACGTGCGGCTGGC	3076
Db	3253	agtggysaaccttgygaacccgtcttccaaagagacttgygaacaaagtgcactcatgttcccg	3312
Qy	3077	TTCCAGGCCAGATGTTTCCAGAAAGTTATCTGTCGTCTTCTGTGTGGGAATGATC	3136

[illegible]

Db 2794 caactcgagagacagctcagcagctcctcctcctatagctgggagacatcagcagcaccac 2853
 QY 2816 AAGAGAAAGGAGGAGGTTTCAGTGCCTGGTCAGATATAGTGTGTACATCCAGCTGCTG 2675
 Db 2854 aacagccaccgagacatcgcaagaagctgagccaaagaagagagacccctggaagctgctg 2913
 QY 2876 GGGGG-----CCCTTGGCATTCCTGGTTATATAGGCCCTTTTTCATGCTCAATGTAGGC 2728
 Db 2914 gaggctgacaagagcgcagcagcaggcaggtcaggtcctcctgctcactgaggactacatgag 2973
 QY 2729 AGCACCGCCTTCACAGCACCTGGTGGTTGAGTTAGTATGATCAAGCAAGAGGCGGAACACC 2788
 Db 2974 gcatctgacctctcactcctcctcctcagcactcctccttccatgctgtaaccaagctgctc 3033
 QY 2789 ACTGTGACT-----CGAGGAGACGACCTCGGTGAGTGAAGCAGCAGTAAAGACAACTCT 2842
 Db 3034 ggcgtggtcctcaactatgctcagcctctgagactgtagcccatcgtaacagcgagct 3093
 QY 2843 CATATGAGTACTATGCCAGATCTAGCCCTCTTCATGGCAGTCATGCTGATCTGAAA 2902
 Db 3094 caagagacacgaagaagtcggtcgtcagcgtcctcgtgagccctgaggacttcaagaaggatc 3153
 QY 2803 GCCATTCGAGG-----AGTGTCTTTTGTCAAGGAGCAGCTGCGAGCTTCCCGGCTG 2956
 Db 3154 ggcgtggttgcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3213
 QY 2957 CATGACGAGCTTTTCCGAGAGATCTTCGAAAGCCCTATGAAGTTTGTGACAGCAGCCCC 3016
 Db 3214 caagtgcagcctgcaacagatcctgctcagcctcctcctcctcctcctcctcctcctcctcctc 3273
 QY 3017 ACAGGAGGATTTTCAACAGGTTTCCAAAGACATGATGATGATGATGATGATGATGATGATG 3076
 Db 3274 agtgggaacctgaggaacccgtctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3333
 QY 3077 TTCCAGGCGGAGATTTTCAACAGGTTTCCGAGGATTTTCCGAGGATTTTCCGAGGATTTTCC 3136
 Db 3334 gaggatcacaagatgctcactgctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3393
 QY 3137 GCAGAGACTTTCCCGTGGTTCCTTGTGGCAGTGGAGCCCTTGTGATCTTCTTTTCACTG 3196
 Db 3394 ctgtgtgcagcagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3453
 QY 3197 CTGACACATTTGTTCCAGGAGTCTGATTCGGAGGAGTGAAGGCTTGGAGACATATACGCGAG 3256
 Db 3454 gtccagaggtctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3513
 QY 3257 TCACCTTTTCTTCCACATCCAGTTCAGATACGAGGAGGCTTGGCAGCAGCTTCCAGCTTAC 3316
 Db 3514 tcccggtctatctccatcttcaacgagaccttgcctggtgcagcgtcattcctcctcctcctc 3573
 QY 3317 AATTAAGGAGGAGGATTTTCCAGATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3376
 Db 3574 gaggagcagagagcgtctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3633
 QY 3377 TTTTGTGTGTAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3436
 Db 3634 tattaaccacagatcgtggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3693
 QY 3437 GCCCTCATCACACACAGGAGGAGTATGATCGTTTATGACAGGAGGAGGAGGAGGAGGAGGAG 3496
 Db 3694 tgcactgctcgttgcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3753
 QY 3497 TATCGGAGTCTCGCCATCTTATGCTGTCCAGATTAACGGGAGCTTCCAGTTTACGGTTC 3556
 Db 3754 ttggtggcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3813
 QY 3557 AGACTGGAGTGTGAGAGAGGAGTGCATTCACCTCGGTGAGGAGGAGGAGGAGGAGGAGGAGG 3616
 Db 3814 cggatgctcactcgtgaagcaac 3873
 QY 3617 AAGACTCTGTCTTGGAGAGCACTGCGAGAAATTAAAGAACAAAGGCTCCCTCCCTGACTGG 3676
 Db 3874 gaggactgagaa---gagagcgccctcgtgcaaatccagagagacagctcgcgccagcagctg 3930

QY 3677 CCCACGAGGAGAGGTGACTTTTGAACACGACGATGAGTACCCAGAAACCTCCT 3736
 Db 3931 cccaagctggtccgagtggaattcccggaactacgtcgtcgtacccgagagagacttgagc 3990
 QY 3737 CTGCTCCAAAGAAAGTATCTTCCATCGATTCAAACCTTAAGAAAGATGAGATTTGTGGGG 3796
 Db 3991 ttcgtctcacaatcaatgatagtcagatcaatggtgggagaaagtgatcgttggtg 4050
 QY 3797 CGACAGGATGAGGAGGAGTCTCCTCGTGGGATGGCCCTTCCGCTGCTGAGTATATCT 3856
 Db 4051 cggacggagagcttggaagctgctcctcctcctcctcctcctcctcctcctcctcctcctcctc 4110
 QY 3857 GGAGGCTCATCAAGATTTGATGAGTGAATCACTGATATTTGGCTTCCGAGCTCCGA 3916
 Db 4111 gaaggagagatcatatcgtatgatacaacacacacacacacacacacacacacacacacacac 4170
 QY 3917 AGCAAACTCTATCATCTCTCAAGAGCGGAGCTTGTGAGTGCAGTGCAGTGCAGTGCAGTGC 3976
 Db 4171 ttcaagatcacaatcctcccaagacccctgttctgttctgtgtcctcctcctcctcctcctcctc 4230
 QY 3977 TTGACCCCTTCAACACAGTACATGAAAGCAGATTTGGATGCCCTGAGAGGACACAC 4036
 Db 4231 ctggaaccattcagccagctactcgtatgagaagatcgtgagctcctcgtgagctgccccac 4290
 QY 4037 ATGAAGAATGATTTGCTACCTTACCTTGAACCTTGAATCTGAATGATGAGATGAG 4096
 Db 4291 ctgaaggactcgtgtcagccctcctcctcctcctcctcctcctcctcctcctcctcctcctc 4350
 QY 4097 GATTAACCTCTGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4156
 Db 4351 gagaacccacagctgtcgtgagcagcagcagctgtgtgtcctcctcctcctcctcctcctcctc 4410
 QY 4157 TGTAGATTTGATTTTATGATGAAGCCACAGCTGCATGACACAGAGACAGACTTATG 4216
 Db 4411 acgaagactcctgtgtgtatgagagccagcagcagcagcagcagcagcagcagcagcagcagc 4470
 QY 4217 ATTGAAGACATTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4276
 Db 4471 atcagatccacatcgtgacacagatcagagactgacagcgtcctcctcctcctcctcctc 4530
 QY 4277 CTGACAGGAGTCTTATGAGGCTCGATGATGATGATGATGATGATGATGATGATGATGATG 4336
 Db 4531 ctcaacacacatcagatgatacaacaaggtgatgtctgtgcaagaaggaatccagagag 4590
 QY 4337 TTTGACACCCCATCGGCTCTTCTG 4360
 Db 4591 tacggtcccatcgagactcctg 4614

Search completed: November 9, 2001, 15:58:48
 Job time: 10446 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 12:19:42 ; Search time 153.99 Seconds
(without alignments)
7128.637 Million cell updates/sec

Title: US-09-528-031-1
Perfect score: 4847
Sequence: 1 GGCTCATGCTCGGAGCGGTG.....AAAAAAAAAGGCGGCCCG 4847

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCRTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4847	100.0	4847	3	US-09-061-400-1 Sequence 1, Appli
2	4775.4	98.5	4781	2	US-09-001-273-1 Sequence 1, Appli
3	4775.4	98.5	4781	4	US-08-843-459A-1 Sequence 1, Appli
4	511.2	10.5	5011	1	US-08-463-092B-3 Sequence 3, Appli
5	511.2	10.5	5011	2	US-08-462-109A-3 Sequence 3, Appli
6	511.2	10.5	5011	2	US-08-460-907B-3 Sequence 3, Appli
7	511.2	10.5	5011	3	US-08-463-179A-3 Sequence 3, Appli
8	511.2	10.5	5011	3	US-08-463-384B-3 Sequence 3, Appli
9	506.4	10.4	5011	1	US-08-141-893-1 Sequence 1, Appli
10	506.4	10.4	5011	1	US-08-463-092B-1 Sequence 1, Appli
11	506.4	10.4	5011	2	US-08-462-109A-1 Sequence 1, Appli
12	506.4	10.4	5011	2	US-08-460-907B-1 Sequence 1, Appli
13	506.4	10.4	5011	3	US-08-463-179A-1 Sequence 1, Appli
14	506.4	10.4	5011	3	US-08-461-384B-1 Sequence 1, Appli
15	506.4	10.4	5011	3	US-08-407-207A-1 Sequence 1, Appli
16	420.4	8.7	5889	1	US-08-463-092B-5 Sequence 5, Appli
17	420.4	8.7	5889	2	US-08-462-109A-5 Sequence 5, Appli
18	420.4	8.7	5889	2	US-08-460-907B-5 Sequence 5, Appli
19	420.4	8.7	5889	3	US-08-463-179A-5 Sequence 5, Appli
20	420.4	8.7	5889	3	US-08-461-384B-5 Sequence 5, Appli
21	370	7.6	4931	4	US-08-726-320-2 Sequence 2, Appli
22	370	7.6	4931	4	US-09-208-716-2 Sequence 2, Appli
23	346	7.1	4877	2	US-08-404-531B-7 Sequence 7, Appli
24	346	7.1	4877	2	US-08-404-531B-7 Sequence 7, Appli
25	346	7.1	4877	3	US-08-476-900A-8 Sequence 8, Appli
26	346	7.1	4877	3	US-08-476-900A-8 Sequence 8, Appli
27	346	7.1	4877	3	US-08-488-546A-7 Sequence 7, Appli

28	346	7.1	4877	3	US-08-488-546A-8	Sequence 8, Appli
29	344.4	7.1	5175	4	US-08-972-927-4	Sequence 4, Appli
30	342.6	7.1	5110	2	US-08-404-531B-4	Sequence 4, Appli
31	342.6	7.1	5110	2	US-08-404-531B-5	Sequence 5, Appli
32	342.6	7.1	5110	3	US-08-476-900A-4	Sequence 4, Appli
33	342.6	7.1	5110	3	US-08-476-900A-5	Sequence 5, Appli
34	342.6	7.1	5110	3	US-08-488-546A-4	Sequence 4, Appli
35	342.6	7.1	5110	3	US-08-488-546A-5	Sequence 5, Appli
36	340.8	7.0	463	2	US-09-001-273-3	Sequence 3, Appli
37	340.8	7.0	463	2	US-09-061-400-3	Sequence 3, Appli
38	340.8	7.0	463	4	US-08-843-459A-3	Sequence 3, Appli
39	336.2	6.9	5232	4	US-08-972-927-1	Sequence 1, Appli
40	330.8	6.8	2454	2	US-08-404-531B-32	Sequence 32, Appli
41	330.8	6.8	2454	3	US-08-476-900A-32	Sequence 32, Appli
42	330.8	6.8	2454	3	US-08-488-546A-32	Sequence 32, Appli
43	294	6.1	2294	2	US-08-404-531B-30	Sequence 30, Appli
44	294	6.1	2294	3	US-08-476-900A-30	Sequence 30, Appli
45	294	6.1	2294	3	US-08-488-546A-30	Sequence 30, Appli

ALIGNMENTS

```
RESULT 1
US-09-061-400-1
: Sequence 1, Application US/09061400
: Patent No. 6077936
:
: GENERAL INFORMATION:
: APPLICANT: SHYUAN, Andrew
: TITLE OF INVENTION: NOVEL MULTIDRUG RESISTANCE ASSOCIATED
: NUMBER OF INVENTIONS: POLYPEPTIDE
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESS: LAHIVE & COCKFIELD, LLP
: STREET: 28 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/061,400
: FILING DATE: 16-APRIL-1998
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Elizabeth A. Hanley
: REGISTRATION NUMBER: 33,505
: REFERENCE/DOCKET NUMBER: MNT-056CP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 742-4214
: INFORMATION FOR SEQ ID NO: 1:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4847 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 116..4426
:
: US-09-061-400-1
:
: Query Match 100.0%; Score 4847; DB 3; Length 4847;
: Best Local Similarity 100.0%; Pred. No. 0;
: Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: 1 GGCTCATGCTCGGAGCGGTGAGCGGCTGCGCGGTTGCTTGAGACGAGGCGCAG 60
: |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

D	b	1	GGCTCAGTCGCGGGAGGGTGGCTTGGACGGGCTGGCCGGCTGTGCTTGAGACAGGGGGCGAG	60
Q	y	61	GAATTCGATGTGAACCTAACAGTCTGTGTGAGCCCTTGAAACTCCACTCAGAGAAAGTGA	120
D	b	61	GAATTCGATGTGAACCTAACAGTCTGTGTGAGCCCTTGAAACTCCACTCAGAGAAAGTGA	120
Q	y	121	GGATATCGACATATGAGAAACAGATATATCATATCCCAAGTCCCTGGGTATAGAAAGTGAGGGA	180
D	b	121	GGATATCGACATATGAGAAACAGATATATCATATCCCAAGTCCCTGGGTATAGAAAGTGAGGGA	180
Q	y	181	GAGAACGACACTTCTGGGAGCGACGACAGAGACCGTGAAGATTTCCAAAGTTTCAGAGAAATCG	240
D	b	181	GAGAACGACACTTCTGGGAGCGACGACAGAGACCGTGAAGATTTCCAAAGTTTCAGAGAAATCG	240
Q	y	241	ACCGTTGAATGCCAAGATGCCCTTGGAAACAGCAGCCGAGCCGAGGCGCTCTCTTTGA	300
D	b	241	ACCGTTGAATGCCAAGATGCCCTTGGAAACAGCAGCCGAGCCGAGGCGCTCTCTTTGA	300
Q	y	301	TGCGTCATGCAATTCACAGCTCGAATTCCTGGATGGAGAGACATCCCAAGGGGAAAGTACA	360
D	b	301	TGCGTCATGCAATTCACAGCTCGAATTCCTGGATGGAGAGACATCCCAAGGGGAAAGTACA	360
Q	y	361	TCATGCGTTAGTGTCTGGAAGCCCATCCGAGACTATTTGCCAAACACACAGCACCAGTGA	420
D	b	361	TCATGCGTTAGTGTCTGGAAGCCCATCCGAGACTATTTGCCAAACACACAGCACCAGTGA	420
Q	y	421	CAATGCTGGGCTTTTTCCTGTATGACTTTTTCGTGTGCTTCTCTCTGCCCCGTGTGGC	480
D	b	421	CAATGCTGGGCTTTTTCCTGTATGACTTTTTCGTGTGCTTCTCTCTGCCCCGTGTGGC	480
Q	y	481	CCACAAGAAAGGGGAGCTCTCAATGGAAGAGCGTGGCTCTGTGCCAAGCAAGCATCTTC	540
D	b	481	CCACAAGAAAGGGGAGCTCTCAATGGAAGAGCGTGGCTCTGTGCCAAGCAAGCATCTTC	540
Q	y	541	TGACGTGACACTGACAGAAAGACTATAGAGACTGTGTGCAAGAAAGCTGAATGAAAGTTGGCC	600
D	b	541	TGACGTGACACTGACAGAAAGACTATAGAGACTGTGTGCAAGAAAGCTGAATGAAAGTTGGCC	600
Q	y	601	AGACGCTGTCTCCCTGCGAAGGGTGTGTGTGATCTTCTGCCGACCAAGCTATATCTGTC	660
D	b	601	AGACGCTGTCTCCCTGCGAAGGGTGTGTGTGATCTTCTGCCGACCAAGCTATATCTGTC	660
Q	y	661	CATGTGTGCTGTATGATCAACGAGGTGGGTGGGCTCATGTGACCAAGGCTTCATATGTGA	720
D	b	661	CATGTGTGCTGTATGATCAACGAGGTGGGTGGGCTCATGTGACCAAGGCTTCATATGTGA	720
Q	y	721	ACACCTTTGGAGATATACCCAGGCAACAGAGTCTTAACCTGCAGTACAGCTTGTGTAGT	780
D	b	721	ACACCTTTGGAGATATACCCAGGCAACAGAGTCTTAACCTGCAGTACAGCTTGTGTAGT	780
Q	y	781	GCTGGGCTCTCTCTCTGACGGAATTCGTGGGCTTGTGTCGCTTGCACCTGACTGTGGGCAAT	840
D	b	781	GCTGGGCTCTCTCTCTGACGGAATTCGTGGGCTTGTGTCGCTTGCACCTGACTGTGGGCAAT	840
Q	y	841	GAATTACCGAACCGGTCCGCTTGGGGGGGGGCATCCTAACCAATGCAATTAAGAAAT	900
D	b	841	GAATTACCGAACCGGTCCGCTTGGGGGGGGGCATCCTAACCAATGCAATTAAGAAAT	900
Q	y	901	CCTTAAGTTAAAGAACATTTAAGAAATCCCGGGTGTGCTATCAACATTGGCTCAA	960
D	b	901	CCTTAAGTTAAAGAACATTTAAGAAATCCCGGGTGTGCTATCAACATTGGCTCAA	960
Q	y	961	CGATGGGACAGATGTTTATGAGCAGACACCGTTGGCAGCCTGTGCTGGAGGAGACCGT	1020
D	b	961	CGATGGGACAGATGTTTATGAGCAGACACCGTTGGCAGCCTGTGCTGGAGGAGACCGT	1020
Q	y	1021	TGTTGCCATCTTAAGCATGATTTATATGTAATTAATCTGGGACCAACAGGCTTCTGGG	1080
D	b	1021	TGTTGCCATCTTAAGCATGATTTATATGTAATTAATCTGGGACCAACAGGCTTCTGGG	1080
Q	y	1081	ATCAGCTGTTTTATCTCTTTTACCAGCAATGATGTTTGATCAGGGCTCACAGCAT	1140
D	b	1081	ATCAGCTGTTTTATCTCTTTTACCAGCAATGATGTTTGATCAGGGCTCACAGCAT	1140

Qy	1141	TTTCAGAGAAAAATGCGTGGCCGCCACCGATGACGTGTCCAGAAAGATGAATGAAGTCT	1200
Dp	1141	TTTTCAGAGAAAAATGCGTGGCCGCCACCGATGAAAGTGTCCAGAAAGATGAATGAAGTCT	1200
Qy	1201	TACTTACATTAAATTTATCAAAATGTATGTCCTGGGTCAAAAGCATTTTCTCAGAGTGTCA	1260
Dp	1201	TACTTACATTAAATTTATCAAAATGTATGTCCTGGGTCAAAAGCATTTTCTCAGAGTGTCA	1260
Qy	1261	GAAAATCCGGAGAGAGAGCGTGGATATTGGAAAAAGCCGGGTACTTCCAGAGCATCAC	1320
Dp	1261	GAAAATCCGGAGAGAGAGCGTGGATATTGGAAAAAGCCGGGTACTTCCAGAGCATCAC	1320
Qy	1321	TGTGGGTGTGGCTCCCAATTTGCTGGTGTATGTGGCACACGGGTGAGCACTTCTCGTTCATAT	1380
Dp	1321	TGTGGGTGTGGCTCCCAATTTGCTGGTGTATGTGGCACACGGGTGAGCACTTCTCGTTCATAT	1380
Qy	1381	GACCTGTGGCTTGATCTGCACGACACAGCGCTTTCACAGTGTGTACAGTCTTCAATTC	1440
Dp	1381	GACCTGTGGCTTGATCTGCACGACACAGCGCTTTCACAGTGTGTACAGTCTTCAATTC	1440
Qy	1441	CATACCTTTGCTTTGAAAGTAAACCGCTTTTCAGTAAAGTCCCTCCTCAGAAAGCTCAGT	1500
Dp	1441	CATACCTTTGCTTTGAAAGTAAACCGCTTTTCAGTAAAGTCCCTCCTCAGAAAGCTCAGT	1500
Qy	1501	GGCTGTGACAGATTTAAAGATTTGTTTCTAATGGAAGGTTCACATGATTAAGAACAA	1560
Dp	1501	GGCTGTGACAGATTTAAAGATTTGTTTCTAATGGAAGGTTCACATGATTAAGAACAA	1560
Qy	1561	ACGACGCAGTCTCACACTCAAAAGTAAAGATGAAAAATGGCACTTGGCATGGGACTCTC	1620
Dp	1561	ACGACGCAGTCTCACACTCAAAAGTAAAGATGAAAAATGGCACTTGGCATGGGACTCTC	1620
Qy	1621	CCACTCCAGATCCAGAACACGCGCCCAAGGTGACCCCCAAAATGAAAAAAGACAAAGAGGC	1680
Dp	1621	CCACTCCAGATCCAGAACACGCGCCCAAGGTGACCCCCAAAATGAAAAAAGACAAAGAGGC	1680
Qy	1681	TTCCAGGGGCAAGAAAGAAAGGTGAGGACGCTGCACGGCCACTGAGCATCAGGGGTCT	1740
Dp	1681	TTCCAGGGGCAAGAAAGAAAGGTGAGGACGCTGCACGGCCACTGAGCATCAGGGGTCT	1740
Qy	1741	GGCAGACAGAAAAAGGCCACCTCTCTGTGCACAGTGCACGAGCGCCAGTCCGGAAGAGA	1800
Dp	1741	GGCAGACAGAAAAAGGCCACCTCTCTGTGCACAGTGCACGAGCGCCAGTCCGGAAGAGA	1800
Qy	1801	AGAAAGCAGACATCCACTGTGGCCACCTGGGCTTACAGAGGACACTGTACACATCTGA	1860
Dp	1801	AGAAAGCAGACATCCACTGTGGCCACCTGGGCTTACAGAGGACACTGTACACATCTGA	1860
Qy	1861	TTCTGGACATCCAAAGAGGCTAAACTGTGTGGAAATCTCGGCACATGTGGGAAGTGGAAAAAC	1920
Dp	1861	TTCTGGACATCCAAAGAGGCTAAACTGTGTGGAAATCTCGGCACATGTGGGAAGTGGAAAAAC	1920
Qy	1921	CTCTCTCATTTACAGCAATTTTAAAGCCAGATGACGCTTCTAGAGGGCAGCATTTGCATCAG	1980
Dp	1921	CTCTCTCATTTACAGCAATTTTAAAGCCAGATGACGCTTCTAGAGGGCAGCATTTGCATCAG	1980
Qy	1981	TGCAACTTGGCTTATGTGTGGCCAGCAGACGCTGTGATCTCTCAATGCTACTGTAGAGACAA	2040
Dp	1981	TGCAACTTGGCTTATGTGTGGCCAGCAGACGCTGTGATCTCTCAATGCTACTGTAGAGACAA	2040
Qy	2041	CATCTCTTTTGGGAAGGAATATATGAAGAAGATCAACTCTGTGTGCAACAGCTGCTG	2100
Dp	2041	CATCTCTTTTGGGAAGGAATATATGAAGAAGATCAACTCTGTGTGCAACAGCTGCTG	2100
Qy	2101	CTGTAGGCTCTGACCTGGCCATTTTCTCCACAGCAGCACTGACGAGATTTGGAGACGAGG	2160
Dp	2101	CTGTAGGCTCTGACCTGGCCATTTTCTCCACAGCAGCACTGACGAGATTTGGAGACGAGG	2160
Qy	2161	AGCCCACTGAGGGGTGGAGAGGCCAGAGATCAGCTTGGCCGGGCTGTATATAGTA	2220
Dp	2161	AGCCCACTGAGGGGTGGAGAGGCCAGAGATCAGCTTGGCCGGGCTGTATATAGTA	2220

QY	2221	CAGAGCATCTCAATCCTTGSAGCACCCCTCAGTGCCTTAAATGCCCATGTGGGCACCA	2280
Dp	2221	CAGAGCATCTCAATCCTTGSAGCACCCCTCAGTGCCTTAAATGCCCATGTGGGCACCA	2280
QY	2281	CATCTTCAATAGTGCCTTCCGGAAACATCTCAAGTCCAAAGACAGTCTGTTGTTACCCA	2340
Dp	2281	CATCTTCAATAGTGCCTTCCGGAAACATCTCAAGTCCAAAGACAGTCTGTTGTTACCCA	2340
QY	2341	CCAGTTACAGTACCTGGTTGCTACGTGATGAAGTGAATCTTTCATGAAGAGGGCTGTATTAC	2400
Dp	2341	CCAGTTACAGTACCTGGTTGCTACGTGATGAAGTGAATCTTTCATGAAGAGGGCTGTATTAC	2400
QY	2401	GGAAGAAGCACCCCATGAGSAGACTGATGAATTTAAATGCTGACTTGTACCATTTTTAA	2460
Dp	2401	GGAAGAAGCACCCCATGAGSAGACTGATGAATTTAAATGCTGACTTGTACCATTTTTAA	2460
QY	2461	TAACTCTGCTGGAGAGACACCGGCAGTTGAGATCAATTCATAAAAAAGAAACAGAGG	2520
Dp	2461	TAACTCTGCTGGAGAGACACCGGCAGTTGAGATCAATTCATAAAAAAGAAACAGAGG	2520
QY	2521	TTCCAGAAAGAGTCACACAGACAAAGGGTCTTAAACAGATCAATTAAGAAAGAAAAAGC	2580
Dp	2521	TTCCAGAAAGAGTCACACAGACAAAGGGTCTTAAACAGATCAATTAAGAAAGAAAAAGC	2580
QY	2581	AGTAAAGCCAGAGAGAGAGGCGACTTGTGCAGCTGGAAAGAAAGGGCAGGTTCACTGCC	2640
Dp	2581	AGTAAAGCCAGAGAGAGAGGCGACTTGTGCAGCTGGAAAGAAAGGGCAGGTTCACTGCC	2640
QY	2641	CTGGCAGTATATGATGCTCTACATCCAGGCTGCTGGGGGCCCTTGGCATTCCTGGTAT	2700
Dp	2641	CTGGCAGTATATGATGCTCTACATCCAGGCTGCTGGGGGCCCTTGGCATTCCTGGTAT	2700
QY	2701	TATGGCCCTTTTCATGCTGTAATGTAGGACAGACCGGCTTCACAGCACCCTGGTGTGAAGTTA	2760
Dp	2701	TATGGCCCTTTTCATGCTGTAATGTAGGACAGACCGGCTTCACAGCACCCTGGTGTGAAGTTA	2760
QY	2761	CTGATCAAGCAGAGAGAGGGGGAACACACCTGTGACTCGAGGAAACAGACCTCGGTAG	2820
Dp	2761	CTGATCAAGCAGAGAGAGGGGGAACACACCTGTGACTCGAGGAAACAGACCTCGGTAG	2820
QY	2821	TGACAGCATGAAGACACATCTCATATGACAGTATGTCGACACATCTACGCCCTCCAT	2880
Dp	2821	TGACAGCATGAAGACACATCTCTATATGACAGTATGTCGACACATCTACGCCCTCCAT	2880
QY	2881	GGCAGTCATGCTGATCTCTGAAAGCCATTCGAGAGATTGCTTTGTCAAGGGCAGCTGCG	2940
Dp	2881	GGCAGTCATGCTGATCTCTGAAAGCCATTCGAGAGATTGCTTTGTCAAGGGCAGCTGCG	2940
QY	2941	AGCTTCTCCCGGCTGCTATGACGAGCTTTTCCGAGAGATCTCTTCAAGCCCTATGAAGTT	3000
Dp	2941	AGCTTCTCCCGGCTGCTATGACGAGCTTTTCCGAGAGATCTCTTCAAGCCCTATGAAGTT	3000
QY	3001	TTTTGACACGACCCCCACAGGAGGATTTCTCAACAGTTTTTCCAAAGACATGAGTGAAGT	3060
Dp	3001	TTTTGACACGACCCCCACAGGAGGATTTCTCAACAGTTTTTCCAAAGACATGAGTGAAGT	3060
QY	3061	TGAGCTGCGGCTGCGCTTCCAGGCGCAGATGTTCATCAGACGTTATCTCGTGTCTCT	3120
Dp	3061	TGAGCTGCGGCTGCGCTTCCAGGCGCAGATGTTCATCAGCAGAACGTTATCTCGTGTCTCT	3120
QY	3121	CTGTGTGGGAATGATCGACGAGACTTCCCGTGGTCTTGTGTGGCAGTGGGGCCCTCTGT	3180
Dp	3121	CTGTGTGGGAATGATCGACGAGACTTCCCGTGGTCTTGTGTGGCAGTGGGGCCCTCTGT	3180
QY	3181	CATCCTCTTTTCAATCCTGTGCACATTTGTCTCCAGGGCTCGAATTCGGGAGCTGAACGCTCT	3240
Dp	3181	CATCCTCTTTTCAATCCTGTGCACATTTGTCTCCAGGGCTCGAATTCGGGAGCTGAACGCTCT	3240
QY	3241	GGACAAATTCACGAGACACCTTTCCTCACCACATCAGTCCAGATATACAGGGCTCTGC	3300
Dp	3241	GGACAAATTCACGAGACACCTTTCCTCACCACATCAGTCCAGATATACAGGGCTCTGC	3300
QY	3301	CACCATCCAGCCCTTCAATTAAGGGCAGAGTTTCTGCACACATACAGAGACTGCTGGA	3360

Db	3301	CACCATCCAGCCCTTACATATAAAGGACGGAGTTCTGCAACAAATACAGAGACTGCTGGA	3360
Qy	3361	TGACAAACCAAGCTCTTTTATTTTGTATTTACGTGCGATGCGGTGGCGCTGCGCGCT	3420
Db	3361	TGACAAACCAAGCTCTTTTATTTTGTATTTACGTGCGATGCGGTGGCGCTGCGCGCT	3420
Qy	3421	GGACCTCATCAGATCGCCCTCATCACACACAGGGGCTGATGATCGTTCTTATGACAGG	3480
Db	3421	GGACCTCATCAGATCGCCCTCATCACACACAGGGGCTGATGATCGTTCTTATGACAGG	3480
Qy	3481	GCAATTTCCCAACGCTTATGCGGGTCTCGCATCTCTTATGCTGTCCAGTTAAACGGGCT	3540
Db	3481	GCAGATTTCCCAACGCTTATGCGGGTCTCGCATCTCTTATGCTGTCCAGTTAAACGGGCT	3540
Qy	3541	GTTCCAGTTTACGGTTCAGACTGGCATCTGAGACAGAAAGTCGATTCACCTCGGTGGAAG	3600
Db	3541	GTTCCAGTTTACGGTTCAGACTGGCATCTGAGACAGAAAGTCGATTCACCTCGGTGGAAG	3600
Qy	3601	GATCAATCACTACATTTAAGACTGTCTCTTGAAGCACCTGCCAGAAATTAAGAACAGGC	3660
Db	3601	GATCAATCACTACATTTAAGACTGTCTCTTGAAGCACCTGCCAGAAATTAAGAACAGGC	3660
Qy	3661	TCCCTCCCTTACCTGCGCCCAAGAGAGAGTGCACCTTTGAGAAACGAGATAGGTA	3720
Db	3661	TCCCTCCCTTACCTGCGCCCAAGAGAGAGTGCACCTTTGAGAAACGAGATAGGTA	3720
Qy	3721	CCGAGAAAACCTCCCTCGCTCTTAAAGAAAGTACTTCACGATCAACCTTAAAGAGAA	3780
Db	3721	CCGAGAAAACCTCCCTCGCTCTTAAAGAAAGTACTTCACGATCAACCTTAAAGAGAA	3780
Qy	3781	GATTGGCAATTGTGGGGGCGACAGATCAGGGAAGTCTCTGTGGGGATGGCCCTCTTCG	3840
Db	3781	GATTGGCAATTGTGGGGGCGACAGATCAGGGAAGTCTCTGTGGGGATGGCCCTCTTCG	3840
Qy	3841	TCTGTGAGAGTTATCTGGAGGCTGCATCAAGATTTGATGAGAGTGAATACGATATGG	3900
Db	3841	TCTGTGAGAGTTATCTGGAGGCTGCATCAAGATTTGATGAGAGTGAATACGATATGG	3900
Qy	3901	CTTTCCGCACTCCGAGCAAAACCTCTATCATTTCTCTCAAGACCGGTCCTTTAGTGG	3960
Db	3901	CTTTCCGCACTCCGAGCAAAACCTCTCTATCATTTCTCTCAAGACCGGTCCTTTAGTGG	3960
Qy	3961	CACCTGCAGATCAAAATTTGAGCCCTTCAACACAGTACACTGGAAGACAGATTTGGATGC	4020
Db	3961	CACCTGCAGATCAAAATTTGAGCCCTTCAACACAGTACACTGGAAGACAGATTTGGATGC	4020
Qy	4021	CCTGAGAGGACACACATGAAAGATTTATGCTCAGCTACTGTGAACTTGAATCTGGA	4080
Db	4021	CCTGAGAGGACACACATGAAAGATTTATGCTCAGCTACTGTGAACTTGAATCTGGA	4080
Qy	4081	AGTATGGAAGAAATGGGATTAATCTTCACGTGGGGGAACGGCAAGCTTGTGATAGCTAG	4140
Db	4081	AGTATGGAAGAAATGGGATTAATCTTCACGTGGGGGAACGGCAAGCTTGTGATAGCTAG	4140
Qy	4141	AGCCCTGCTCCGCCACGTGAAGATTTGATTTTATGATGAAAGCACACAGTCCCATGGACAC	4200
Db	4141	AGCCCTGCTCCGCCACGTGAAGATTTGATTTTATGATGAAAGCACACAGTCCCATGGACAC	4200
Qy	4201	AGAGACAGACTTATTTGATTCAGAGACCATCCGAGAAGATTTGGCAGACTGTACATGCT	4260
Db	4201	AGAGACAGACTTATTTGATTCAGAGACCATCCGAGAAGATTTGGCAGACTGTACATGCT	4260
Qy	4261	GACCATTTCCCATGGCTCGACACAGGGTTTAAAGCTCCGATAGGATTAATGCTGCGCCA	4320
Db	4261	GACCATTTCCCATGGCTCGACACAGGGTTTAAAGCTCCGATAGGATTAATGCTGCGCCA	4320
Qy	4321	GGGACAGGTGTGTGAGTTTGAACCCCATGGTCTCTTGTCTCCACAGACAGTTCCCGATT	4380
Db	4321	GGGACAGGTGTGTGAGTTTGAACCCCATGGTCTCTTGTCTCCACAGACAGTTCCCGATT	4380
Qy	4381	CTATCCATGTTTCTGCTGCGAGAAACAAGTGTGCTCAAGGGCTGACTCTCCCTGT	4440
Db	4381	CTATCCATGTTTCTGCTGCGAGAAACAAGTGTGCTCAAGGGCTGACTCTCCCTGT	4440

```

Db 4381 CTATGCCATGTTGCTGCTGCAGAGAACAGAGTCGTCAAGGGCTGCTCTCCCTGT 4440
OY 4441 TGAGAGAGTCTCTTTCTTTAGAGATGCGATCCCTGCTGGGGGGGGCCCTTCATC 4500
Db 4441 TGAGAGAGTCTCTTTCTTTAGAGATGCGATCCCTGCTGGGGGGGGCCCTTCATC 4500
OY 4501 GCGTCCCTCCAGAACCTTGCCTTTCGATTTTATCTTTCGACAGAGATTCGGAT 4560
Db 4501 GCGTCCCTCCAGAACCTTGCCTTTCGATTTTATCTTTCGACAGAGATTCGGAT 4560
OY 4561 TGGCTGTGTCTTACCTTTAGAGAGATGATATTGATTTATTTATTCATAT 4620
Db 4561 TGGCTGTGTCTTACCTTTAGAGAGATGATATTGATTTATTTATTCATAT 4620
OY 4621 TCAGTAAACAAATTTAGTTTGTCTTATTAATGACCTTAAGAGTTCAGGAACTG 4680
Db 4621 TCAGTAAACAAATTTAGTTTGTCTTATTAATGACCTTAAGAGTTCAGGAACTG 4680
OY 4681 TATATATTTGATCAGAGGCTATTAATGAAGCTTTATAGCTATATCTATATAT 4740
Db 4681 TATATATTTGATCAGAGGCTATTAATGAAGCTTTATAGCTATATCTATATAT 4740
OY 4741 AATCTGTACATAGCTTATTTACAGTGAAGAAATGATAGCTGTATTTATTTAAAT 4800
Db 4741 AATCTGTACATAGCTTATTTACAGTGAAGAAATGATAGCTGTATTTATTTAAAT 4800
OY 4801 AAGACTGTCTTAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 4847
Db 4801 AAGACTGTCTTAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 4847

RESULT 2
US-09-001-273-1
: Sequence 1, Application US/09001273
: Patient No. 5994130
: GENERAL INFORMATION:
: APPLICANT: SHYAN, Andrew
: TITLE OF INVENTION: NOVEL MULTIDRUG RESISTANCE-ASSOCIATED
: TYPE OF INVENTION: POLYPEPTIDE
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Testa, Hurwitz & Thibault
: STREET: 125 High St.
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/001,273
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: FENTON, Gillian M
: REGISTRATION NUMBER: 36,508
: REFERENCE/DOCKET NUMBER: MTL-001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 248-7000
: TELEFAX: (617) 248-7100
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4781 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 2..4360
: US-09-001-273-1

```

```

Query Match 98.5%: Score 4775.4; DB 2; Length 4781;
Best Local Similarity 99.8%: Pred. No. 0;
Matches 4773; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 67 TGATGTGAACATACAGTCTGTGAGCCCTTGAAACCTCCACTCAGAGAAATGAAGATAT 126
Db 1 TGATGTGAACATACAGTCTGTGAGCCCTTGAAACCTCCACTCAGAGAAATGAAGATAT 60

OY 127 GCACATAGAAAAGATATATCATCCCAAGTCCCTGGGATATAGAACTGAGAGAGAAAC 186
Db 61 GCACATAGAAAAGATATATCATCCCAAGTCCCTGGGATATAGAACTGAGAGAGAAAC 120

OY 187 CAGCACTTCTGGGAGCGCAGAGACCGTGAAGATTCAGATTCCAGAGAACCTGCACCTT 246
Db 121 CAGCACTTCTGGGAGCGCAGAGACCGTGAAGATTCAGATTCCAGAGAACCTGCACCTT 180

OY 247 GGAATGCCAAGATGCTTGAAGAACAGACCCGAGCCGAGGGCTCTCTGTGATGCTC 306
Db 181 GGAATGCCAAGATGCTTGAAGAACAGACCCGAGCCGAGGGCTCTCTGTGATGCTC 240

OY 307 CATGATTTCTAGCTCAGATCAATTCCTGGATGAGAGATCCCAAGGAAAGTACCATCATG 366
Db 241 CATGATTTCTAGCTCAGATCAATTCCTGGATGAGAGATCCCAAGGAAAGTACCATCATG 300

OY 367 CTGAGTCTCTGAAGCCCATCCGACTACTTGCAGAACACAGACCCAGTGGACATATG 426
Db 301 CTGAGTCTCTGAAGCCCATCCGACTACTTGCAGAACACAGACCCAGTGGACATATG 360

OY 427 TGGGCTTTTTCCTGTATGACTTTTTCGTGCTTCTTCTGCGCCGTGTGGCCACAA 486
Db 361 TGGGCTTTTTCCTGTATGACTTTTTCGTGCTTCTTCTGCGCCGTGTGGCCACAA 420

OY 487 GAAGGGAGAGCTCTCAATAGAAAGAGCTGTGCTCTGTCAGAGACAGAGTCTTGACAGT 546
Db 421 GAAGGGAGAGCTCTCAATAGAAAGAGCTGTGCTCTGTCAGAGACAGAGTCTTGACAGT 480

OY 547 GAAGTGCAGAGACTAGAGAGAGACTGTGCAAGAGAGTGAATGAATGGAGCCAGAGCC 606
Db 481 GAAGTGCAGAGACTAGAGAGAGACTGTGCAAGAGAGTGAATGAATGGAGCCAGAGCC 540

OY 607 TGTCTCCCTGCGAAGGGTGTGTGTGATCTTCTGCCGACACAGCTCATCTGTCCATGT 666
Db 541 TGTCTCCCTGCGAAGGGTGTGTGTGATCTTCTGCCGACACAGCTCATCTGTCCATGT 600

OY 667 GTGCTGATGATCAGCAGCTGGCTGGCTCAGTGTGAGACAGGCTTCATGTGAACACCT 726
Db 601 GTGCTGATGATCAGCAGCTGGCTGGCTCAGTGTGAGACAGGCTTCATGTGAACACCT 660

OY 727 CTGAGATATACCCAGGCAAGAGTCTAACCTGCAGTACAGCTTGTGTAGTCTGGG 786
Db 661 CTGAGATATACCCAGGCAAGAGTCTAACCTGCAGTACAGCTTGTGTAGTCTGGG 720

OY 787 CTTCTCTCTGACGAAATCGTGGCTCTTGGTCTGCTTGACACTGACTTGGGCAATTGAATA 846
Db 721 CTTCTCTCTGACGAAATCGTGGCTCTTGGTCTGCTTGACACTGACTTGGGCAATTGAATA 780

OY 847 CCGAAGCGGTGCGCTTGGCGGGGGCCATCCATACATATGGCATTTAGAGATCTTAA 906
Db 781 CCGAAGCGGTGCGCTTGGCGGGGGCCATCCATACATATGGCATTTAGAGATCTTAA 840

OY 907 GTTAAGAACATTTAAAGAAATCCCTGGGTGAGCTCATCAACATTTCTCCAAAGATG 966
Db 841 GTTAAGAACATTTAAAGAAATCCCTGGGTGAGCTCATCAACATTTCTCCAAAGATG 900

OY 967 GCAGAGATGTTTGAAGCAGACCGCTTGGCAGCTGTGCTGGTGAAGACCCGTTGTTGC 1026
Db 901 GCAGAGATGTTTGAAGCAGACCGCTTGGCAGCTGTGCTGGTGAAGACCCGTTGTTGC 960

OY 1027 CATCTTAGCATGATTTTAAATGTAATATCTGGAGACCAAGGCTCTGGGATGACG 1086
Db 961 CATCTTAGCATGATTTTAAATGTAATATCTGGAGACCAAGGCTCTGGGATGACG 1020

```

Qy	1087	TGTTTTATCTCTTTTATCCACCAATGATGTTTGCATCAGGGCTCACAGCATATTTGAC	1146
Dp	1021	TGTTTTATCTCTTTTATCCACCAATGATGTTTGCATCAGGGCTCACAGCATATTTGAC	1080
Qy	1147	GAGAAATGCGTGGCCGCCACCGGATGACGTGCGCCAGGAAGTGAAGTTCCTACTTA	1206
Dp	1081	GAGAAATGCGTGGCCGCCACCGGATGACGTGCGCCAGGAAGTGAAGTTCCTACTTA	1140
Qy	1207	CATTAAATTTATCAAAATGTATGCTGGGTCAAAAGCATTTCACAGAGTTCACAGAAAT	1266
Dp	1141	CATTAAATTTATCAAAATGTATGCTGGGTCAAAAGCATTTCACAGAGTTCACAGAAAT	1200
Qy	1267	CCGGGAGGAGGAGCGTCGGAATATGGAAAAACCGGGTACTCCAGACATCACGTGGG	1326
Dp	1201	CCGGGAGGAGGAGCGTCGGAATATGGAAAAACCGGGTACTCCAGACATCACGTGGG	1260
Qy	1327	TGTGGCTCCCATTTGTGTGTGTGATTTGCCACGCGTGGAGACCTTCCTGTTCATATGACCT	1386
Dp	1261	TGTGGCTCCCATTTGTGTGTGTGATTTGCCACGCGTGGAGACCTTCCTGTTCATATGACCT	1320
Qy	1387	GGGCTTGATCTGCACGACGACACAGGCTTTCACAGTGGTACAGTCTTCAATTCATGAC	1446
Dp	1321	GGGCTTGATCTGCACGACGACACAGGCTTTCACAGTGGTACAGTCTTCAATTCATGAC	1380
Qy	1447	TTTTGCTTTGAAAGTAAACCGCTTTTCAGTAAAGTCCCTCAGAAAGCTCAGTGGCTGT	1506
Dp	1381	TTTTGCTTTGAAAGTAAACCGCTTTTCAGTAAAGTCCCTCAGAAAGCTCAGTGGCTGT	1440
Qy	1507	TGACAGATTTAAGAGTGTGTTCTTAATGGAAGAGTTCACATGATTAAGAAACAAACGAC	1566
Dp	1441	TGACAGATTTAAGAGTGTGTTCTTAATGGAAGAGTTCACATGATTAAGAAACAAACGAC	1500
Qy	1567	CAGCGCCACATCAAGATAGATAGATGAAAAATGCGACCTTGGATGGGACTCCTCCACATC	1626
Dp	1501	CAGCGCCACATCAAGATAGATAGATGAAAAATGCGACCTTGGATGGGACTCCTCCACATC	1560
Qy	1627	CAGTATCCAGAACTCGCCCAAGCTGACCCCCCAAAATGAAAAAAGACAAGAGGCTTCAG	1686
Dp	1561	CAGTATCCAGAACTCGCCCAAGCTGACCCCCCAAAATGAAAAAAGACAAGAGGCTTCAG	1620
Qy	1687	GGGCAAAAGAGAGAGGTGAGGATGCGAGGCGCATGAGCATCAGGCGGTCTGGCAGA	1746
Dp	1621	GGGCAAAAGAGAGAGGTGAGGATGCGAGGCGCATGAGCATCAGGCGGTCTGGCAGA	1680
Qy	1747	GCAGAAAGGCGCACCTCCTCTGACAGTAGACGACGCGCCCACTCCCGAAGAGAAAGG	1806
Dp	1681	GCAGAAAGGCGCACCTCCTCTGACAGTAGACGACGCGCCCACTCCCGAAGAGAAAGG	1740
Qy	1807	CAGACACATCACCTGGGCCACTGCGCTTTACAGAGGACACTGCACAGCATCATCTGGA	1866
Dp	1741	CAGACACATCACCTGGGCCACTGCGCTTTACAGAGGACACTGCACAGCATCATCTGGA	1800
Qy	1867	GATCCAGAGAGGTAAATGGTTGGATTCGCGGCAAGTGGGAACATGGAAAAAACCCTCT	1926
Dp	1801	GATCCAGAGAGGTAAATGGTTGGATTCGCGGCAAGTGGGAACATGGAAAAAACCCTCT	1860
Qy	1927	CATTTCAGCAATTTAGCCAGATGACGCTTCAGAGGCGACACATTCGATCAGTGAAC	1986
Dp	1861	CATTTCAGCAATTTAGCCAGATGACGCTTCAGAGGCGACACATTCGATCAGTGAAC	1920
Qy	1987	CTTGCGTTATGTGGCCACGACAGGCTTGGATCCTCAATGCTACTCTGAGAGACAACTCT	2046
Dp	1921	CTTGCGTTATGTGGCCACGACAGGCTTGGATCCTCAATGCTACTCTGAGAGACAACTCT	1980
Qy	2047	GTTTGGGAAGGAATATGATGAAAGAAAGTAAACAATCTGTGCTGAACAGCTGCTGCTAG	2106
Dp	1981	GTTTGGGAAGGAATATGATGAAAGAAAGTAAACAATCTGTGCTGAACAGCTGCTGCTAG	2040
Qy	2107	GCTGACCTGGCCATTTCTTCCACGACGACGACTGACGGAATGGAGAGCAGAGGCCAA	2166
Dp	2041	GCTGACCTGGCCATTTCTTCCACGACGACGACTGACGGAATGGAGAGCAGAGGCCAA	2100

QY	2167	CCTAGCGGTTGGCAGCGCCAGAGGATCAAGCTTGGCCGGGCTGTATAGTGACAGAG	2228
Db	2101	CCTAGGGGGTGGGAGGCGCCAGAGGATCAAGCCTTGGCCGGGCTGTATAGTGACAGAG	2160
QY	2227	CATGTACATCCTGGAGCAACCCCTCAGTGGCTTATAGTATGGCCATGAGGCAACATATTT	2286
Db	2161	CATGTACATCCTGGAGCAACCCCTCAGTGGCTTATAGTATGGCCATGAGGCAACATATTT	2220
QY	2287	CAATAGTGTATCCGGAAACATCTCAAGTCCAAAGACAGTTCGTTGTTTACCCACAGTT	2346
Db	2221	CAATAGTGTATCCGGAAACATCTCAAGTCCAAAGACAGTTCGTTGTTTACCCACAGTT	2280
QY	2347	ACAGTACCTGGTTGACGTGTATGAAGTGATCTTCATGAAGAAGGCGCTGTATTCGGAAG	2406
Db	2281	ACAGTACCTGGTTGACGTGTATGAAGTGATCTTCATGAAGAAGGCGCTGTATTCGGAAG	2340
QY	2407	AGGCACCCATGAGGAACCTGTATGAATTTAAATGTTGCTATGTACATTTTAAATAACT	2466
Db	2341	AGGCACCCATGAGGAACCTGTATGAATTTAAATGTTGCTATGTACATTTTAAATAACT	2400
QY	2467	GTTCCTGGAGAGACACCGCCAGTTGAGATCAATTTAAAAAGAAACAGTGGTTACA	2586
Db	2401	GTTCCTGGAGAGACACCGCCAGTTGAGATCAATTTAAAAAGAAACAGTGGTTACA	2460
QY	2527	GAAGAAGTCACAAAGCAAGGGTCTTAAACAGGATCAATTAAGAAAGAAAAAGCATATA	2586
Db	2461	GAAGAAGTCACAAAGCAAGGGTCTTAAACAGGATCAATTAAGAAAGAAAAAGCATATA	2520
QY	2587	GCCAGAGGAAGGCGAGCTTGTGCAGCTGGAAGAAAGGCGCAGGGTTCAGTCCCTGGTC	2646
Db	2521	GCCAGAGGAAGGCGAGCTTGTGCAGCTGGAAGAAAGGCGCAGGGTTCAGTCCCTGGTC	2580
QY	2647	AGTATATGTTCTACATTCGACGCTCGTGGGGCCCTTGGCATTCGTTATATATATGC	2706
Db	2581	AGTATATGTTCTACATTCGACGCTCGTGGGGCCCTTGGCATTCGTTATATATATGC	2640
QY	2707	CCTTTCATVCTGATGTATGAGCGACACCGCCTTCAGCACCTGTGTGTAGTTACTGGAT	2766
Db	2641	CCTTTCATVCTGATGTATGAGCGACACCGCCTTCAGCACCTGTGTGTAGTTACTGGAT	2700
QY	2767	CAACGAAGGAAGGGGGAACCACTGATCGTAGGGAAGGAGACCTGGTGAGTGAAG	2826
Db	2701	CAACGAAGGAAGGGGGAACCACTGATCGTAGGGAAGGAGACCTGGTGAGTGAAG	2760
QY	2827	CATGAAGGACATCTCATATATGACAGTACTATGCACACATCTACGCCCTCCATGGCAGT	2886
Db	2761	CATGAAGGACATCTCATATATGACAGTACTATGCACACATCTACGCCCTCCATGGCAGT	2820
QY	2887	CATGCTGATCTGAAAGCCATTTGAGGAGTTGCTTTTGTCAAGGGCAGCGTCGAGCTTC	2946
Db	2821	CATGCTGATCTGAAAGCCATTTGAGGAGTTGCTTTTGTCAAGGGCAGCGTCGAGCTTC	2880
QY	2947	CTCCCGGCTGCATGACGAGCTTTTCCGAGAGATCTTTCGAAGCCTATGAACTTTTGTGA	3006
Db	2881	CTCCCGGCTGCATGACGAGCTTTTCCGAGAGATCTTTCGAAGCCTATGAACTTTTGTGA	2940
QY	3007	CACGACCCCAACAGGGAGATTCGAAAGGTTTCCAAAGACATGAGATGAAGTGAAGT	3066
Db	2941	CACGACCCCAACAGGGAGATTCGAAAGGTTTCCAAAGACATGAGATGAAGTGAAGT	3000
QY	3067	GGGCGTGGCTTCCAGGCGCAAGATGTTTCATCTCAAGACGTTATCTGTGTCTTCTGTGT	3126
Db	3001	GGGCGTGGCTTCCAGGCGCAAGATGTTTCATCTCAAGACGTTATCTGTGTCTTCTGTGT	3060
QY	3127	GGGAATGATGCGAGAGTCTTCCCGTGGTTCTTGTGGCAGTGGGGCCCTTGTCTATCTT	3186
Db	3061	GGGAATGATGCGAGAGTCTTCCCGTGGTTCTTGTGGCAGTGGGGCCCTTGTCTATCTT	3120
QY	3187	CTTTTCAGTCTCTGCAACTTCTTCGAGGGGCTCGATGTTGGGAGCGTGAAGCGTCTGACAA	3246
Db	3121	CTTTTCAGTCTCTGCAACTTCTTCGAGGGGCTCGATGTTGGGAGCGTGAAGCGTCTGACAA	3180
QY	3247	TATACGCGAGTCACCTTTCCTCTCCACATCAAGTCCACACATACAGGGCGCTTGCACCAT	3306

Db 3181 TATACGACGACACCTTCTCTCCACATCAGCTCCAGCATACAGGCGCTTGCACCAT 3240
 QY 3307 CCACGCTACATAAAGGCGAGAGTTCTGCAGAGATACAGAGAGCTGTGATGACAA 3366
 Db 3241 CCACGCCCTACAAATAAAGGCGAGAGTTCTGCAGAGATACAGAGAGCTGTGATGACAA 3300
 QY 3367 CCAAGCTCTTTTTTTTTTTTGTAGTGTGCGATGCGGTGCTGTGCGGTGCGACCT 3426
 Db 3301 CCAAGCTCTTTTTTTTTTTTGTAGTGTGCGATGCGGTGCTGTGCGGTGCGACCT 3360
 QY 3427 CATAGACATGCGCCGATACACACAGGCGGTATGATGCTGTATGACAGCGGACAT 3486
 Db 3361 CATAGACATGCGCCGATACACACAGGCGGTATGATGCTGTATGACAGCGGACAT 3420
 QY 3487 TCCCCAGCCTATGCGGCTGCGGCATCTTTATGCTGTCAGTTAAAGGCGGTGCTCA 3546
 Db 3421 TCCCCAGCCTATGCGGCTGCGGCATCTTTATGCTGTCAGTTAAAGGCGGTGCTCA 3480
 QY 3547 GTTACGCTCAGACTGCGATGTGAGACAGAAAGCTGATTCACCTGCGGTGAGAGATCAA 3606
 Db 3481 GTTACGCTCAGACTGCGATGTGAGACAGAAAGCTGATTCACCTGCGGTGAGAGATCAA 3540
 QY 3607 TCACATACATTAAGCTGCTGCTGGAAGCAGCTGCGCAATTAAGAAAGGCTCCCTC 3666
 Db 3541 TCACATACATTAAGCTGCTGCTGGAAGCAGCTGCGCAATTAAGAAAGGCTCCCTC 3600
 QY 3667 CCGTACGCTGCGCCAGAGAGAGAGGTGACCTTTGAGAACGAGAGATGAGTACCGAGA 3726
 Db 3601 CCGTACGCTGCGCCAGAGAGAGAGGTGACCTTTGAGAACGAGAGATGAGTACCGAGA 3660
 QY 3727 AAACCTCCCT 3786
 Db 3661 AAACCTCCCT 3720
 QY 3787 CATGTGGGGGGGAGAGATCAGGGAAGTCTCTGCTGGGGATGGGCGCTCCGCTGCTGCT 3846
 Db 3721 CATGTGGGGGGGAGAGATCAGGGAAGTCTCTGCTGGGGATGGGCGCTCCGCTGCTGCT 3780
 QY 3847 GGAAGTTATCTGAGAGCTGCATCAAGATTTGATGAGTGAAGTCAAGTATTTGGCTTGC 3906
 Db 3781 GGAAGTTATCTGAGAGCTGCATCAAGATTTGATGAGTGAAGTATTTGGCTTGC 3840
 QY 3907 CGACCTCGAAGCAAACTCTCTATCATCTCTCAAGAGCGGCTGCTCTCACTGGCAGCTGT 3966
 Db 3841 CGACCTCGAAGCAAACTCTCTATCATCTCTCAAGAGCGGCTGCTCTCACTGGCAGCTGT 3900
 QY 3967 CAGATCAAAATTTGGACCCCTCAACACAGTACAGTGAAGATTTGGATGGCCCTGGA 4026
 Db 3901 CAGATCAAAATTTGGACCCCTCAACACAGTACAGTGAAGATTTGGATGGCCCTGGA 3960
 QY 4027 GAGACACACATGAAGAAATTTGCTCAGCTACCTCTGAACCTTGAATCTGAAGTAT 4086
 Db 3961 GAGACACACATGAAGAAATTTGCTCAGCTACCTCTGAACCTTGAATCTGAAGTAT 4020
 QY 4087 GGAGAAATGGGATTAATCTCTAGTGGGGGAACGAGCTCTTGTGACATAGAGCCCT 4146
 Db 4021 GGAGAAATGGGATTAATCTCTAGTGGGGGAACGAGCTCTTGTGACATAGAGCCCT 4080
 QY 4147 GCTCCGACACGTAAAGATTTGATTTTATGATGAAGCCACACACTGCGATGACACAGAGAC 4206
 Db 4081 GCTCCGACACGTAAAGATTTGATTTTATGATGAAGCCACACACTGCGATGACACAGAGAC 4140
 QY 4207 AGACTTATTTGATTAAGAGACCATCCGAAAGATTTGACAGCTGTACATGCTGACCAT 4266
 Db 4141 AGACTTATTTGATTAAGAGACCATCCGAAAGATTTGACAGCTGTACATGCTGACCAT 4200
 QY 4267 TGCCCATGCGCTGACACAGGTTCTAGGCTCCGATAGATTTAGTGTGCGGCGCCAGAGACA 4326
 Db 4201 TGCCCATGCGCTGACACAGGTTCTAGGCTCCGATAGATTTAGTGTGCGGCGCCAGAGACA 4260
 QY 4327 GGTTGGAGTTTACACACCCATGCTGCTCTGTCACAAAGACAGTCCCATTTCTATGC 4386

Db 4261 GGTGTGAGTTTACACACCCATGCTCTCTGTGTCACAGACAGTTCGGATCTATGC 4320
 QY 4387 CATGTTGCTGCTGAGAGAAAGTGTGCTGACAGGCTGCTCTCTGTTGACGA 4446
 Db 4321 CATGTTGCTGCTGAGAGAAAGTGTGCTGACAGGCTGCTCTCTGTTGACGA 4380
 QY 4447 AGTCTCTTCTGTTAGACATTCGCTCCCTGCGGGGGGGCCCTTATGCGGCTC 4506
 Db 4381 AGTCTCTTCTGTTAGACATTCGCTCCCTGCGGGGGGGCCCTTATGCGGCTC 4440
 QY 4507 TCCATCCGAACCTTGCCTTCTGATTTTATCTTTGTCACAGAGATTCGGATTTGCTT 4566
 Db 4441 TCCATCCGAACCTTGCCTTCTGATTTTATCTTTGTCACAGAGATTCGGATTTGCTT 4500
 QY 4567 GTGTGTTCACTTTTAGGAGAGATCATATTTGATTTATTTATTTATTCATTTATCT 4626
 Db 4501 GTGTGTTCACTTTTAGGAGAGATCATATTTGATTTATTTATTTATTCATTTATCT 4560
 QY 4627 AAACAAATTTAGTTTGTCTTATTTGCACTGTAAGAGTTCAAGGAACGTTATTTAT 4686
 Db 4561 AAACAAATTTAGTTTGTCTTATTTGCACTGTAAGAGTTCAAGGAACGTTATTTAT 4620
 QY 4687 AATGTATCAGAGCGCTTAAATGAAGCTTTTATGCTGATATCTATTTATTTATCT 4746
 Db 4621 AATGTATCAGAGCGCTTAAATGAAGCTTTTATGCTGATATCTATTTATTTATCT 4680
 QY 4747 GTACATACCTATATTTTACAGTGAAGATTAAGCTTTTATTTATTTATTTATTTAT 4806
 Db 4681 GTACATACCTATATTTTACAGTGAAGATTAAGCTTTTATTTATTTATTTATTTATTT 4740
 QY 4807 TGTGCTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 4847
 Db 4741 TGTGCTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 4781

RESULT 3
 US-08-843-459A-1
 : Sequence 1, Application US/08843459A
 : Patent No. 6162616
 : GENERAL INFORMATION:
 : APPLICANT: SHYJAN, Andrew
 : TITLE OF INVENTION: NOVEL MULTIDRUG RESISTANCE-ASSOCIATED
 : TITLE OF INVENTION: POLYPEPTIDE
 : NUMBER OF SEQUENCES: 8
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: LAHIVE & COCKFIELD, LLP
 : STREET: 28 State Street
 : CITY: Boston
 : STATE: Massachusetts
 : COUNTRY: USA
 : ZIP: 02109
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentln Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/843,459A
 : FILING DATE: 16-Apr-1997
 : CLASSIFICATION: 536
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Hanley, Elizabeth A.
 : REGISTRATION NUMBER: 33,505
 : TELECOMMUNICATION INFORMATION:
 : REFERENCE/DOCKET NUMBER: MNI-056 (formerly MIL-001)
 : TELEPHONE: (617)227-7400
 : TELEFAX: (617)742-4214
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 4781 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear

FEATURE:
NAME/KEY: CDS
LOCATION: 2.4360
US-08-843-459A-1

Query Match 98.5%; Score 4775.4; DB 4; Length 4781;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 4773; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 67 TGAATGAAACATAACAGTGTGTAGCCCTGGAACCTCACTCAGAGAGATGAAGATAT 126
DB 1 TGAATGAAACATAACAGTGTGTAGCCCTGGAACCTCACTCAGAGAGATGAAGATAT 60
OY 127 CCACATAGGAAAAGATATATATATCCCACTCTGGGTATAGAGTGTAGAGAGAAAC 186
DB 61 CCACATAGGAAAAGATATATATATCCCACTCTGGGTATAGAGTGTAGAGAGAAAC 120
OY 187 CAGCACTTCTGGAGGACACAGAGACCGTGAAGATCCAGATTCCAGAGAACTCGACCTT 246
DB 121 CAGCACTTCTGGAGGACACAGAGACCGTGAAGATCCAGATTCCAGAGAACTCGACCTT 180
OY 247 GGAATGCCAAGATGCCCTTGGAAACAGACCCGAGCCGAGGCTCTCTCTTGATGCTTC 306
DB 181 GGAATGCCAAGATGCCCTTGGAAACAGACCCGAGCCGAGGCTCTCTCTTGATGCTTC 240
OY 307 CATGCATTTCTACGCTCAGATTCCTGGATGAGAGCATCCCAAGGAAATACCATCATGG 366
DB 241 CATGCATTTCTACGCTCAGATTCCTGGATGAGAGCATCCCAAGGAAATACCATCATGG 300
OY 367 CTGGATGCTCGAAGCCCATCCGAGCTACTTGCAAACACGACCCCACTGGACAAATGC 426
DB 301 CTGGATGCTCGAAGCCCATCCGAGCTACTTGCAAACACGACCCCACTGGACAAATGC 360
OY 427 TGGGCTTTTCTCTGTATGACTTTTCTGTGCTTCTCTGTGCTGCTGCTGCTGCTGCTGCT 486
DB 361 TGGGCTTTTCTCTGTATGACTTTTCTGTGCTTCTCTGTGCTGCTGCTGCTGCTGCTGCT 420
OY 487 GAAAGGGAGCTCTCAATGGAAGAGCTGTGTCTCTGTCCAAAGCAGAGTCTTGTGAGCT 546
DB 421 GAAAGGGAGCTCTCAATGGAAGAGCTGTGTCTCTGTCCAAAGCAGAGTCTTGTGAGCT 480
OY 547 GAACTGCAAGACTAGAGAGACTGTGCAAGAGAGCTGAATGAATGGGGCAGACGC 606
DB 481 GAACTGCAAGACTAGAGAGACTGTGCAAGAGAGCTGAATGAATGGGGCAGACGC 540
OY 607 TGCCTTCCCTGCGAAGGTTGTGTGATCTTCTGCGCACACAGGCTCATCTGTCAATCGT 666
DB 541 TGCCTTCCCTGCGAAGGTTGTGTGATCTTCTGCGCACACAGGCTCATCTGTCAATCGT 600
OY 667 GTGCTGATGATCAGCAGCTGGCTGCTTCACTGAGTGAACAGCTTCTATGTAACAGCT 726
DB 601 GTGCTGATGATCAGCAGCTGGCTGCTTCACTGAGTGAACAGCTTCTATGTAACAGCT 660
OY 727 CTGGAGTATCCACAGGCAACAGAGTCTAACTGCACTCAGCTTGTGTGTGTGTGTGTGT 786
DB 661 CTGGAGTATCCACAGGCAACAGAGTCTAACTGCACTCAGCTTGTGTGTGTGTGTGTGT 720
OY 787 CCTCTCTCTGACGGAATGCTGGGTCTGTGCTGTGCTGTGCACTGACTTGGGACTTGAATTA 846
DB 721 CCTCTCTCTGACGGAATGCTGGGTCTGTGCTGTGCTGTGCACTGACTTGGGACTTGAATTA 780
OY 847 CCGAACCGGTGTCCTGTCGGGGGGCCATCTTAACATGGGCACTTTAAAGATCTTTAA 906
DB 781 CCGAACCGGTGTCCTGTCGGGGGGCCATCTTAACATGGGCACTTTAAAGATCTTTAA 840
OY 907 GTTAAAGACATTTAAAGAGAAATCCCTGGGTGAGCTCATCAATTTGTCTCAACAGATG 966
DB 841 GTTAAAGACATTTAAAGAGAAATCCCTGGGTGAGCTCATCAATTTGTCTCAACAGATG 900
OY 967 GCAGAGAAATGTTGAGGACAGACCGCTTGGACGCTGTGCTGAGAGACCGGTGTTGC 1026
DB 901 GCAGAGAAATGTTGAGGACAGACCGCTTGGACGCTGTGCTGAGAGACCGGTGTTGC 960

OY 1027 CACCTTAGGCATGATTTATATATGTAATTTATCTGGGACCAACAGGCTTCTGGATCAGC 1086
DB 961 CACCTTAGGCATGATTTATATATGTAATTTATCTGGGACCAACAGGCTTCTGGATCAGC 1020
OY 1087 TGTATTTATCTCTTTTATCCAGCAATGATGTTTGCATCAGGCTTCCACAGATTTTCAG 1146
DB 1021 TGTATTTATCTCTTTTATCCAGCAATGATGTTTGCATCAGGCTTCCACAGATTTTCAG 1080
OY 1147 GAGAAAATCGTGGCCGCCAGGATGAACTGTCCAGAAATGAAATGAAATGAAATGAAAT 1206
DB 1081 GAGAAAATCGTGGCCGCCAGGATGAACTGTCCAGAAATGAAATGAAATGAAATGAAAT 1140
OY 1207 CATTAATTTATCAAAATATATGCTTGGGTCAAAAGATTTTCTCAGAGTGTCAAGAAAT 1266
DB 1141 CATTAATTTATCAAAATATATGCTTGGGTCAAAAGATTTTCTCAGAGTGTCAAGAAAT 1200
OY 1267 CCGCGAGGAGAGCGTCGATATTTGGAAGAAAGCCGGCTACTTCCAGAGCATCAGTGGG 1326
DB 1201 CCGCGAGGAGAGCGTCGATATTTGGAAGAAAGCCGGGTACTTCCAGAGCATCAGTGGG 1260
OY 1327 TGTGCTCCATTTGTGTGTGATTTCCAGCGTGTGACCTTCTCTGTATATGACCTT 1386
DB 1261 TGTGCTCCATTTGTGTGTGATTTCCAGCGTGTGACCTTCTCTGTATATGACCTT 1320
OY 1387 GGGCTTCATCTGACAGCAGCAGAGCTTTCACAGTGTGACAGCTTCAATTTCCATGAC 1446
DB 1321 GGGCTTCATCTGACAGCAGCAGAGCTTTCACAGTGTGACAGCTTCAATTTCCATGAC 1380
OY 1447 TTTTGTCTTGAAGTAAACCGCTTTCAGTAAAGTCCCTCAGAAAGCTCAGTGGCTGT 1506
DB 1381 TTTTGTCTTGAAGTAAACCGCTTTCAGTAAAGTCCCTCAGAAAGCTCAGTGGCTGT 1440
OY 1507 TCAAGATTTAAGAGTTTGTCTTAATGGAAGGTTTCAATGATTAAGAAACAAACACAC 1566
DB 1441 TCAAGATTTAAGAGTTTGTCTTAATGGAAGGTTTCAATGATTAAGAAACAAACACAC 1500
OY 1567 CAGTCTCAGATTCAGATGAGATGAAAAATGCACTTGGCATGGAGTCTCTCCACATC 1626
DB 1501 CAGTCTCAGATTCAGATGAGATGAAAAATGCACTTGGCATGGAGTCTCTCCACATC 1560
OY 1627 CAGTATCCAGAACTGCGCAAGCTGACCCCAAAATGAAAAAGCAAGAGGCTTCCAG 1686
DB 1561 CAGTATCCAGAACTGCGCAAGCTGACCCCAAAATGAAAAAGCAAGAGGCTTCCAG 1620
OY 1687 GGGCAAGAAAGAGAGGTGAGGCACTGACGCGACCTGAGCATCAGGCGTGTGCGAGA 1746
DB 1621 GGGCAAGAAAGAGAGGTGAGGCACTGACGCGACCTGAGCATCAGGCGTGTGCGAGA 1680
OY 1747 GCAGAAAGGCACTCTCTCTGAGACGTACAGCGGCCCAAGTCCCGAAGAGAAAGAG 1806
DB 1681 GCAGAAAGGCACTCTCTCTGAGACGTACAGCGGCCCAAGTCCCGAAGAGAAAGAG 1740
OY 1807 CAAGCATATCCACTGGGCGACCTGCTTACAGAGAGCACTGCAAGATGATGATGGA 1866
DB 1741 CAAGCATATCCACTGGGCGACCTGCTTACAGAGAGCACTGCAAGATGATGATGGA 1800
OY 1867 GATCCAAAGGTTAACTGTTGGAATCTGCGGCACTGTGGAAGTGAAGAAACCTCTCT 1926
DB 1801 GATCCAAAGGTTAACTGTTGGAATCTGCGGCACTGTGGAAGTGAAGAAACCTCTCT 1860
OY 1927 CATTTACAGCAATTTTATGAGCAATGACGCTTCTAAGGGGACAGATGCAATGAGTGAAC 1986
DB 1861 CATTTACAGCAATTTTATGAGCAATGACGCTTCTAAGGGGACAGATGCAATGAGTGAAC 1920
OY 1987 CTTCGCTTATGTGGCCACAGGCTGATCTCAATGCTACTCTGAGAGCAACATCTCT 2046
DB 1921 CTTCGCTTATGTGGCCACAGGCTGATCTCAATGCTACTCTGAGAGCAACATCTCT 1980
OY 2047 GTTGGGAAGAAATATGATGAAGAAAGATACAACTGTGTGTGAACAGTGTCTGCTGAG 2106
DB 1981 GTTGGGAAGAAATATGATGAAGAAAGATACAACTGTGTGTGAACAGTGTCTGCTGAG 2040

QY 2107 GCGTACCTGGCATTTCTCCACAGCAGACCTGACGAGATTGGAGNCGAGGCCAA 2166
Db 2041 GCGTACCTGGCATTTCTCCACAGCAGACCTGACGAGATTGGAGNCGAGGCCAA 2100
QY 2167 CCTAGCGGTGGGAGCGCCAGAGATAGCCTTGCCGGGCTTTGATAGTACAGAG 2226
Db 2101 CCTAGCGGTGGGAGCGCCAGAGATAGCCTTGCCGGGCTTTGATAGTACAGAG 2160
QY 2227 CATCTACCTCTGAGCAGACCCCTCAGTGCCTTAGANTGCCATGTGGCAACCATCTT 2286
Db 2161 CATCTACCTCTGAGCAGACCCCTCAGTGCCTTAGANTGCCATGTGGCAACCATCTT 2220
QY 2287 CAATAGTCTATCCGGGAAACATCTCAAGTCCAAAGACAGTCTGTTTGTACCCACATTT 2346
Db 2221 CAATAGTCTATCCGGGAAACATCTCAAGTCCAAAGACAGTCTGTTTGTACCCACATTT 2280
QY 2347 ACAGTACCTGTTGACTGATGTAAGTATCTTCATGAAAGGGCTGTATTACGGAAG 2406
Db 2281 ACAGTACCTGTTGACTGATGTAAGTATCTTCATGAAAGGGCTGTATTACGGAAG 2340
QY 2407 AGGCACCCATGAGGAATGATGAATTTAAATGATGACTATGCTACATTTTAAATACCT 2466
Db 2341 AGGCACCCATGAGGAATGATGAATTTAAATGATGACTATGCTACATTTTAAATACCT 2400
QY 2467 GTTCTGGGAGAGACACCCGACGTTGAGATGAATTCAAATTAAGAAAGAACACAGTGTACA 2526
Db 2401 GTTCTGGGAGAGACACCCGACGTTGAGATGAATTCAAATTAAGAAAGAACACAGTGTACA 2460
QY 2527 GAAGAAGTCAACAACAAGAGTCTTAAACAGATCAATTAAGAAAGAAAGACAGTAAA 2586
Db 2461 GAAGAAGTCAACAACAAGAGTCTTAAACAGATCAATTAAGAAAGAAAGACAGTAAA 2520
QY 2587 GCCAAGGAGAGGAGCTTTGTGACGTGGAAGAGAAAGGAGGGTTGAGCCCTGCTC 2646
Db 2521 GCCAAGGAGAGGAGCTTTGTGACGTGGAAGAGAAAGGAGGGTTGAGCCCTGCTC 2580
QY 2647 AGTATATGTTCTTACATCCACAGGCTGCTGGGGGCCCTTGCGATTCCTGTTATATAGC 2706
Db 2581 AGTATATGTTCTTACATCCACAGGCTGCTGGGGGCCCTTGCGATTCCTGTTATATAGC 2640
QY 2707 CCTTTTCATGCTGATGATGAGCAGACCGCCTTCAGACCTGATGAGTACTGAT 2766
Db 2641 CCTTTTCATGCTGATGATGAGCAGACCGCCTTCAGACCTGATGAGTACTGAT 2700
QY 2767 CAAGCAAGGAGCGGGAAACACACTGTGACTGAGGGAAACAGACCTGCTGATGACAG 2826
Db 2701 CAAGCAAGGAGCGGGAAACACACTGTGACTGAGGGAAACAGACCTGCTGATGACAG 2760
QY 2827 CATGAAGACAAATCTCTATATGACAGTACTATGCCAGATCTACGCCCTCTCCATGCGAGT 2886
Db 2761 CATGAAGACAAATCTCTATATGACAGTACTATGCCAGATCTACGCCCTCTCCATGCGAGT 2820
QY 2887 CATGCTGATCTGAAAGCATTCGAGAGTGTCTTTTCAAGGGCAGCTGCGGAGCTC 2946
Db 2821 CATGCTGATCTGAAAGCATTCGAGAGTGTCTTTTCAAGGGCAGCTGCGGAGCTC 2880
QY 2947 CTCGCGGTGTCATGACGAGCTTTTCCGAAGGATCTTCGAAGCCCTATGAAGTTTGTGA 3006
Db 2881 CTCGCGGTGTCATGACGAGCTTTTCCGAAGGATCTTCGAAGCCCTATGAAGTTTGTGA 2940
QY 3007 CACGACCCCCACAGGAGGATTTCTACAGGTTTCCAAAGACATGATGAAGTTGACGT 3066
Db 2941 CACGACCCCCACAGGAGGATTTCTACAGGTTTCCAAAGACATGATGAAGTTGACGT 3000
QY 3067 GCGGCTGCGCTTCAGAGCGGAGATGTTATCAGAAAGTATCCGCGGTCTCTCTGCT 3126
Db 3001 GCGGCTGCGCTTCAGAGCGGAGATGTTATCAGAAAGTATCCGCGGTCTCTCTGCT 3060
QY 3127 GGGATATGTCGACAGAGTCTTCCGTTGGTCTCTTGTGGCAGTGGGGCCCTTGTCTATCT 3186
Db 3061 GGGATATGTCGACAGAGTCTTCCGTTGGTCTCTTGTGGCAGTGGGGCCCTTGTCTATCT 3120
QY 3187 CTTTTCAGTCTCTGCACATTTGCTCCAGGGTCTCTGATTCGGGAGCTGAAGCGCTGAGCAA 3246

Db 3121 CTTTTCAGTCTCTGCACATTTGCTCCAGGGTCTCTGATTCGGGAGCTGAAGCGCTTGGACAA 3180
QY 3247 TATCAGCAGTACACCTTCTCTCTCCACATCAGCTCCAGATACAGGGCCCTTGGCACCAT 3306
Db 3181 TATCAGCAGTACACCTTCTCTCTCCACATCAGCTCCAGATACAGGGCCCTTGGCACCAT 3240
QY 3307 CCAGCCTTACAAATAAGGGCAGAGAGTTTCTGACAGATACAGAGAGCTGCTGATGACAA 3366
Db 3241 CCAGCCTTACAAATAAGGGCAGAGAGTTTCTGACAGATACAGAGAGCTGCTGATGACAA 3300
QY 3367 CCAAGCTCTTTTGTGTTTACGTGTGCGATGGGAGGGGCTGGGCTGTGGCGTGGACCT 3426
Db 3301 CCAAGCTCTTTTGTGTTTACGTGTGCGATGGGAGGGGCTGGGCTGTGGCGTGGACCT 3360
QY 3427 CATCAGATCGCCCTCATACACACACAGGGGCTGATGATGCTTTATGCAAGGGCAGAT 3486
Db 3361 CATCAGATCGCCCTCATACACACACAGGGGCTGATGATGCTTTATGCAAGGGCAGAT 3420
QY 3487 TCCCGACGCTATGCGGGGTCTCGCATCTCTTATGCTGTCCAGTTAACGGGCTGTCCA 3546
Db 3421 TCCCGACGCTATGCGGGGTCTCGCATCTCTTATGCTGTCCAGTTAACGGGCTGTCCA 3480
QY 3547 GTTAAAGGTCAGACTGGCATCTGAGAGACAGAGCTGATTCACCTCGGTGGAGAGATCAA 3606
Db 3481 GTTAAAGGTCAGACTGGCATCTGAGAGACAGAGCTGATTCACCTCGGTGGAGAGATCAA 3540
QY 3607 TCACATATTAAGACTCTGCTCTTGGAGACACCTGCCGAATTAAGAACAGGCTCCCTC 3666
Db 3541 TCACATATTAAGACTCTGCTCTTGGAGACACCTGCCGAATTAAGAACAGGCTCCCTC 3600
QY 3667 CCCTGACTGCCCCAGAGAGGAGAGTACCTTTGACAGCAGAGATGAGTACCGAGA 3726
Db 3601 CCCTGACTGCCCCAGAGAGGAGAGTACCTTTGACAGCAGAGATGAGTACCGAGA 3660
QY 3727 AAACCTCCCTCGTCCCTAAAGAAATCTCTCAAGATCAACCTTAAGAAAGATGG 3786
Db 3661 AAACCTCCCTCGTCCCTAAAGAAATCTCTCAAGATCAACCTTAAGAAAGATGG 3720
QY 3787 CATTGTGGGGGAGCAGAGATCAGGAGTCTCGCTGGGATGGGCCCTTCCTCGTCTGCT 3846
Db 3721 CATTGTGGGGGAGCAGAGATCAGGAGTCTCGCTGGGATGGGCCCTTCCTCGTCTGCT 3780
QY 3847 GAGATTATCTGAGGCTGATCAACAGATGATGAGAGTGAATACAGTATTTGGCTTGC 3906
Db 3781 GAGATTATCTGAGGCTGATCAACAGATGATGAGAGTGAATACAGTATTTGGCTTGC 3840
QY 3907 CGACCTCCGAGCAAACTCTATCATCTCTCAAGAGCGGGTGTCACTGGGCACTGT 3966
Db 3841 CGACCTCCGAGCAAACTCTATCATCTCTCAAGAGCGGGTGTCACTGGGCACTGT 3900
QY 3967 CAGATCAATTTTGGACCCCTTCAACACAGTACAGTGAAGACAGATTTGGATGGCCCTGA 4026
Db 3901 CAGATCAATTTTGGACCCCTTCAACACAGTACAGTGAAGACAGATTTGGATGGCCCTGA 3960
QY 4027 GAGGACACACATGAAGAATGATGCTACAGTACCTCTGAATCTGAATGATGAT 4086
Db 3961 GAGGACACACATGAAGAATGATGCTACAGTACCTCTGAATCTGAATGATGAT 4020
QY 4087 GGAAGATGGGATTAATCTCTAGTGGGGAACGAGAGCTCTTGTGCAATGATAGACCCCT 4146
Db 4021 GGAAGATGGGATTAATCTCTAGTGGGGAACGAGAGCTCTTGTGCAATGATAGACCCCT 4080
QY 4147 GCTCCGACACGTAAAGATTTGATTTGATGAAGCCACAGCTGCCATGACAGAGAGC 4206
Db 4081 GCTCCGACACGTAAAGATTTGATTTGATGAAGCCACAGCTGCCATGACAGAGAGC 4140
QY 4207 AGACTATTTGATTAAGAGACCATTCGAGAAACATTTGACAGTGTACATGTGACCAT 4266
Db 4141 AGACTATTTGATTAAGAGACCATTCGAGAAACATTTGACAGTGTACATGTGACCAT 4200
QY 4267 TGCCCATGCGCTGCACACAGGTTCTAGGCTCCGATAGAGATTATGCTGCTGGCCAGGAGCA 4326

Db	4201	TGCCATTGGCGCTGGACACGGTTCTTAGGCTCCGATTAGAGTATTAGTGTGTGGCCACGGAGCA	42620
Qy	4327	GGTGGTGAGATTGACACACCCCATCGGTCTCTTGTGCCAAGCAGATGCCGATTTCTATGC	43866
Db	4261	GGTGGTGAGATTGACACCCCATCGGTCTCTTGTGCCAAGCAGATGCCGATTTCTATGC	4320
Qy	4387	CATGTTTGTCTGTGCAGAGAACAAAGTCCGTGTCAAGGCTGACTTCCTCTTTGACGA	44466
Db	4321	CATGTTTGTCTGTGCAGAGAACAAAGTCCGTGTCAAGGCTGACTTCCTCTTTGACGA	4380
Qy	4447	AGTCTCTTTTCTTTAGAGCATTTGCCATTTCCGTCGCTGGGGGGGGCCCTTATCGCTGC	45066
Db	4381	AGTCTCTTTTCTTTAGAGCATTTGCCATTTCCGTCGCTGGGGGGGGCCCTTATCGCTGC	4440
Qy	4507	TCTACCGAAGACCTGGCTTCTGCATTTTATCTTTTCGACAGCAGTTCCGGATGGGCTT	45666
Db	4441	TCTACCGAAGACCTGGCTTCTGCATTTTATCTTTTCGACAGCAGTTCCGGATGGGCTT	45000
Qy	4567	GTCGTGTTCACTTTTAGGGAGAGTCATATTTTGATTTATGTTATTTATTCATATTCAATG	46266
Db	4501	GTCGTGTTCACTTTTAGGGAGAGTCATATTTTGATTTATGTTATTTATTCATATTCAATG	45600
Qy	4627	AAACAAATTTAGTTTTGTCTTAATTGCACCTCTAAAAGTTCCAGGACCGTTATAT	46866
Db	4561	AAACAAATTTAGTTTTGTCTTAATTGCACCTCTAAAAGTTCCAGGACCGTTATAT	46200

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/141,893
 FILING DATE: 26-OCT-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/407,207
 FILING DATE: 20-MAR-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Steeg, Carol Miernicki
 REGISTRATION NUMBER: 39,539
 REFERENCE/DOCKET NUMBER: Q1546
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (613) 545-2342
 TELEFAX: (613) 545-6853
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5011 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 196..4788
 US-08-463-092B-3

```

1 RESULT 4
2 US-08-463-092B-3
3 Sequence 3, Application US/08463092B
4 Patent No. 576880
5 GENERAL INFORMATION:
6 APPLICANT: Cole, Susan P.C.
7 APPLICANT: Deeley, Roger G.
8 TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
9 TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
10 NUMBER OF SEQUENCES: 9
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: PARQO RESEARCH & DEVELOPMENT INNOVATIONS
13 STREET: Queen's University at Kingston
14 CITY: Kingston
15 STATE: Ontario
16 COUNTRY: CANADA
17 ZIP: K7L 3N6
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: ASCII text
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/08/463,092B
25 FILING DATE: 05-JUN-1995
26 CLASSIFICATION: 435
27 PRIORITY APPLICATION DATA:
28 APPLICATION NUMBER: 07/966,923
29 FILING DATE: 27-OCT-1992
30 CLASSIFICATION: 435
31 PRIORITY APPLICATION DATA:
32 APPLICATION NUMBER: 08/029,340
33 FILING DATE: 8-MAR-1993
34 CLASSIFICATION: 435
35

```

Query Match	10.5%	Score 511.2	DB 1	Length 5011
Best Local Similarity	52.5%	Pred. No. 1.4e-122		
Matches 1367	Conservative	0	Mismatches 1178	Indels 59
			Gaps	9
QY 1811	CACATCCACCTGGGGCCACCTGCGCTTTACAGAGGACACTGCACAGCATCGATCTGGAGATC	1870		
Db 2142	CACATGCACTGGGGCCA--GGAGGACACCTCCACACATGAAGGACATCACTTCTCCATC	2199		
QY 1871	CAAGAGGGTAAACGTGGTTGGGATCTGGCGGCACTGTGGGAATGGAAAAACCTCTCTCAT	1930		
Db 2200	CCCGAAGGTGCTTGTGTGGCGCGTGGTGGGCCACAGTGGGCGTGGGAAAGTCGTCCTGTGCT	2259		
QY 1931	TCAGCCATTTTAAAGGCCACATACCGCTTTGTAGAGGGGCACATTTGCATCAATGTAAGAACTTC	1990		
Db 2260	TCAGGCCCTCTTGGCGTGNATGACAAATGGAGGGGCAAGTGGCTATCAAGGGGCTCCGGT	2319		
QY 1991	GCTATGTGGCCCAAGCAGCGCTGGATCTCTCAATGCTACTGTGAGACAACTCTGT	2050		
Db 2320	GCTATGTGGCCCAAGCAGCGCTGGATCTCTCAATGATGTCTCTTCTCCGAAAAACATCTCTTTT	2379		
QY 2051	GGGAAGGAATATGATGAAGAAAGATACAACTGTGCTGAACAGCTGCGCTTGAGGGCT	2110		
Db 2380	GGATGTCACTGGAGGAACCAATTTACAGTCCGATATACAGGCTGTGTCCTCTCCCA	2439		
QY 2111	GACCTGGCCATTTCTTCCAGAGCAGCACCCTGACGAGATTTGAGAGCCAGAGCAACCTG	2170		
Db 2440	GACCTGGAAATCTGCTCCACAGTGGGGATGAGACAGAGATTTGGCGAAGAAGGGCTGAACCTG	2499		
QY 2171	AGCGGTGGGCAAGCCCAAGAGATACGCTTGGCCCGGCGCTGTATATGATACAGAGCATC	2230		
Db 2500	TCTGGGGGACAGAACACACAGCGGTGAGCCTGCGCGCGCGCTTACTCCAAACCTGCAT	2559		
QY 2231	TACATCTCGGAGAGACCCCTCAGTACCTTTAGATGGCCATGTTGGGCAACACATCTTTCAT	2290		
Db 2560	TACCTCTTCGATGATCCCTCTCAGCACTGGATGGATGGCCATGTGGAAAAACACATCTTTGAA	2619		
QY 2291	AGTGTCTAT-----CCGGAACATCTCAATGCCAAGACAGTTCTGTGTTTATCCACCGAG	2344		
Db 2620	AATGTGATTTGGCCCCAAGGGGATGCTGGAAGAACAGACGCGGATCTTGGTCAACGACAGC	2679		
QY 2345	TTACAGTCTCGGTGTGATGTGATGAAGTCACTTCAATGAAGAGGGCGTGTTCAGGAA	2404		
Db 2660	ATGAGCTACTTGGCCGAGGTGACGCTCATCTGTCATGATGATGAGTGGCGGCAAGATCTCTGGAG	2739		
QY 2405	AGAGGCACCCATGAGGAACGTGATGAATTTAAATGGTGACTATGACATTTTAAATAAC	2464		

Correspondence Address:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,109A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQI-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5011 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 196..4788
US-08-462-109A-3

Query Match 10.5%; Score 511.2; DB 2; Length 5011;
Best Local Similarity 52.5%; Pred. No. 1.4e-122;
Matches 1367; Conservative 0; Mismatches 1178; Indels 59; Gaps 9;

1811 CACATCCACCTGGGCGCAGCTGGCTTACAGAGACACCTGACATCGATCTGGAGATC 1870
Db 2142 CACATTCACCTGGGCGCA--GGAGCGACCTCCACACTGAATGGCACTTCCATCCATC 2199
Qy 1871 CAGAGAGGTAACCTGGTGGATCTGCGGAGAGTGGGAAGTGAACCACTCTCATTT 1930
Db 2200 CCGGAGAGTGGCTTGGTGGCGCTGGTGGGCGCAGGTGGGCTGGGGAAGTGGTGGCTG 2259
Qy 1931 TCAGCATTTTAGGCGCAGATGACGCTTCTAGAGGGGCGAGCATTTGCAATGATGAACTTC 1990
Db 2260 TCAGCCCTTGGCTGAGATGACAAAGTGAAGGGGCGAGCTGGCTATCAAGGGCTCGGTG 2319
Qy 1991 GCTTATGTGGCCAGCAGAGCGCTGATCTCAATGCTACTCTGAGAGACAACATCTGTTT 2050
Db 2320 GCTTATGTGCCACAGAGCGCTGGATTCAGAAATGATTTCTCCGAGAAAAACATCCCTTTT 2379
Qy 2051 GGAAGAGATATGATGAGAAGATACACTGCTGCTGAACAGCTGCTGAGGCGCT 2110
Db 2380 GGAATGACACTGGAGAGAACATATTTACAGGTCCGTGATACAGGCTGTGGCTCTCCCA 2439
Qy 2111 GACCTGGCACTTCTCCAGCAGCAGCACTGAGAGATTTGAGAGCGAGGCCAACCTG 2170
Db 2440 GACCTGGAATCTCTGCCAGTGGGGATGGAGACAGATTTGGGAGAAAGGCGGTGAACCTG 2499

2171 AGCGTGGGCGAGCGCCAGAGATCAAGCCCTTGGCCGGGCGCTGTATATGACAGAGACATC 2230
Db 2500 TCTGGGGGACAGAAAGCAGCGCTGAGCCTGGCCGGGCGCTACTCCACACCTGCATTT 2559
Qy 2231 TACATCTGAGACACCCCTCAGTGGCTTAGATGCCATGTGGGCAACCACTTTCAAT 2290
Db 2560 TACCTCTTGATGATTCCTCCCTCAGAGTGGATGCCCATGTGGGAAACACATCTTTGAA 2619
Qy 2291 AGTGTAT-----CCGAAACATCTCAAGTCCAAAGACAGTCTGTGTTTACCAACAG 2344
Db 2620 AATGTGATTTGGCCCCCAAGGGGATGTGGAAGAACAGCCGGATCTGTGGTCCGACAC 2679
Qy 2345 TTACAGTACCTGGTTGACTGATGATGAAGTATCTCATGAAGAGGGCGTATTACGAA 2404
Db 2680 ATGAGCTACTTCCCGCAGGTGAGCGTCAATATGTCATGATGAGTGGCGGCAATCTGTG 2739
Qy 2405 AGAGGACCCATGAGAGATGATGATTTAATTTAGTACTATGCTACATTTTAAATAC 2464
Db 2740 ATGGGCTCTACAGAGAGTGTGGTGGTGGAGAGCGGCGCTTGCTGATGTTCCGTGCTAC 2799
Qy 2465 CTGTGCTGGAGAGACACCCCACTT---GAGATCAATTTCAAAAAAGAACCAAGTGT 2521
Db 2800 TATGCCAGCAGAGCAGAGCAGAGATGCAAGAGAACGGGGGTACGGGCTGCACGGGT 2859
Qy 2522 TCACAGAAAGATCACAAAGCAAG-----GTCTAAACAGGATCAATTAAGAGG 2573
Db 2860 CCAAGGAGAAAGCAAAATGAGAAATGGCAATGCGATGCTGTGACGACAGTCAAGGGAAG 2919
Qy 2574 AAAAAGCAATAAAGCCAGAGAAAGGACCTTGTGAGCTG----- 2615
Db 2920 CAAGTGCAGAGAGAGCTCAGCAGCTCTCTCTCTATATAGTGGGACATCAGAGGACACAC 2979
Qy 2616 AAGAGAAAGGCGAGGTTAGATGCCCTGGTCAATATATGTGTCTACATCCAGGCTGCTG 2675
Db 2980 AACAGCACCGCAGAACTGAGAAAGCTGAGGCCAAGAGAGAGACCTGGAGCTGATG 3039
Qy 2676 GGGGG-----CCCTTGGCATTCCTGTTATTTATGAGCCCTTTTCACTGCTGAATGTAAG 2728
Db 3040 GAGGCTGACAGAGCGCAGAGAGGGGAGTCAAGCTTCCGTACTGGGACTACATGAAG 3099
Qy 2729 AGCACCCTTACAGCACTGTGGTGTGAGTTACTGTATCAAGCAAGAAAGCGGAAACACC 2788
Db 3100 GCCATCGAATCTTCACTGCTCTCCCTCAGCATCTCTTTCATGTGTAACATGTGTCC 3159
Qy 2789 ACTGTGACT-----CGAGGAAACAGACCTGGTGTGATGACAGCTGAAAGCAATCT 2842
Db 3160 GGGCTGGCTTCCAACTATTTGGCTCAACCTCTGATGATGATGATGCCCATCTCAACGGGACT 3219
Qy 2843 CATATGCACTATATGCTCAAGCATCTACGCGCTCTCCATGGAGTGCATGGTGAATCTGA 2902
Db 3220 CAGAGACACAGAAATCCGGCTGAGCGTATATGAGCCCTGGGCAATTTCAAGAGGATC 3279
Qy 2903 GCCATTCGAGG-----AGTTGCTTTTCAAGGGGACGCTGGAGCTTCTCCCGGCTG 2956
Db 3280 GCGGTGTTGGCTACATCAATGGCGGTGTCCATGGGGGAGATCTGGCTTCCGCTGTGCG 3339
Qy 2957 CATGACAGACTTTTCCGAAGATCTTCGAAGCCCTATGAAGTTTGAACAGACACCC 3016
Db 3340 CACGTGACCTGTGTCACAGCATCTGCGGTACCATGATGATCTTTGAAGCGAGACCCC 3399
Qy 3017 ACAGGAGATTTCTCAACAGGTTTCCAAAGACATGATGAAGTGAACGTCGGCGCTGGC 3076
Db 3400 AGTGGGAACCTGGTGAACCGCTTCTCAAGAGAGCTGGAACACATGATGATCCCG 3459
Qy 3077 TTCAGAGCCGAGATTTTCAAGAGGTTATCTGATGCTGTTCTTCTGTGGGAATGATC 3136
Db 3460 GAGGTATCAAGATGTTTCAATGGGCTCCCTGTTCAAGTCAATGTCCTGATCGTTATC 3519
Qy 3137 GCAGAGTCTTCCCGTGGTTCTTGTGGCAGTGGGCGCCCTGTGATCCTCTTTTCAATC 3196
Db 3520 CTGCTGGCCACGCGCAATCCGCGCATCATCAATCCCGCTTGGCTCATTCCTCTTC 3579
Qy 3197 CTGCACTTGTCTCCAGGGTCTGATTTGGGAGCTTAAGCGTGTGACAAATATACAGCAG 3256

1811 CACATCCACTGGGCCACCTGGCTTACAGAGACATGCACGATCGATCGAGATC 1870
1111 1111111111 111 111111 11111 11111
2142 CACATTCACCTGGGCCA--GGAGCCACCCTCCACACTGAAATGCATCCTTCTCCATC 2199
1871 CAGAGGGTAAACTGGTGGATCTCGGACAGTGGGAAAGTGAACCACTCTCTCAT 1930
1111 1111 1111 111 1111 1111 1111 1111
2200 CCGGAAAGTGGCTTGGTGGCCGTGGGGCCAGGTGGGTGGGAAAGTCGCTCTGCTC 2259
1931 TCAGCATTTTAAAGCAATGACGCTTCTAGAGGCGACATTGCATCACTGAAACCTTC 1990
1111 1111 1111 1111 1111 1111 1111 1111
2260 TCAGCCCTTGGCTGGATGGAATGGAAGTGGAGGGGACAGTGGCTATCAAGGCTCCG 2319
1991 GCTTATGGGCCACAGCGCCGATCCGATGCTACTGCTAGAGACACATCCGTTT 2050
1111 1111 1111 1111 1111 1111 1111 1111
2320 GCTATGTGCCACAGCAGACGCTGGATTCAGATGATTTCTCTCGGAAACATCTTTT 2379
2051 GGGAAAGATATGATGAAAGAAATACAATCTGTGTGAACAGCTGCTCCCTGAGGCT 2110
1111 1111 1111 111 1111 1111 111 1111
2380 GGATGTACGTGGAGGAACCATATTAAGTCCGTGATACAGGCTGTGCTCTCCCA 2439
2111 GACCTGGCATTTCTCCAGCAGCAGCAGCTGACGAGATTGGAGAGCGAGACCACTG 2170
1111 1111 1111 111 1111 1111 111 1111
2440 GACCTGAAATCTTCGCCAGTGGGATCGACAGATTTGGCGAAGGCGGTGAACCTG 2499
2171 AGCGTGGGACAGCGCAGAGATCAGCCTTGGCCGCGCTGTATAGTACAGAGATC 2230
1111 1111 111 111 1111 1111 1111 1111
2500 TCTGGGGACAGACAGACAGCTGAGCCTGGCCGCGGTACTCCACGCTGCAT 2559
2231 TACATCTGGACAGACCCCTCAGTCCCTTGAATGCCATGTGGGACACCATCTTCAT 2290
1111 1111 1111 111 1111 1111 1111 1111
2560 TACCTTCGATGATCCCTCTCAGCAGTGAATGCCATGTGGAAACACATCTTTGAA 2619
2231 AGTGTAT-----CCGAAACATCTCAAGTCCAGACAGTTCGTTTTCACACAG 2344
1111 1111 1111 111 1111 1111 1111 1111
2620 AATGATTTGGCCCCAAGGGGATGCTGAAGAACAGCGGATTTGGTCCAGCAGC 2679
2345 TTACAGTACCTGGTGTAGCTGTGATGAAGTATCTCATGAAGAGGCGTGTATACGAA 2404
1111 1111 111 111 1111 1111 111 1111
2680 ATGACCTCTTCCCGCAGGTGAGCTCATCATGATGATGAGGGGCAAGATCTCTGAG 2739
2405 AGAGCACCCTAGAGAACTGATGAATTTAAATGCTAGTATGCTAACATTTTAAATAC 2464
1111 1111 1111 111 1111 111 111 1111
2740 ATGGCTCTTCCACAGACTGCTGCTCGAGAGCGGCTTGCAGTATCTCTGCTAC 2799
2465 CTGTTGCTGGGAGACACCGCCAGT---GAGATCAATTCAAAAAGAAACAGCTGT 2521
1111 1111 111 111 1111 111 111 1111
2800 TATGCCAGACAGCAGCAGAGAGATGCGAGAGAGAGGGGTCCAGGCTCAGCGGT 2859
2522 TCACAGAGAACTCACAGACAAAG-----GTCTTAAACAGAGATCAATTAAGAAAG 2573
1111 1111 111 111 111 111 111 1111
2860 CCAGGGAAGGAAGCAAGCAATGGAATGAGCATGCTGTGACGAGACAGTGCAGGAAAG 2919
2574 AAAAAGCATTAAGCCAGAGAGAGGAGCTTGTGCAGCTGG----- 2615
1111 1111 111 111 111 111 111 1111
2920 CAACGTCAGAGACACTAGCTAGCTCTCTCTATAGTGGGACATCAGACGACACAC 2979
2616 AAGAGAAAGGAGGGTTCAGTGGCCCTGTGCTAGTATATGCTGTACATCCAGGCTGCTG 2675
1111 1111 111 111 111 111 111 1111
2980 AACACACCGCAGAACTGAGAAAGCTGAGGCCAAGAGAGAGAGCTCGAAGCTGATG 3039
2676 GGGGC-----CCCTTGGCATTTCTGTGTTATATAGCCCTTTTATGCTGATGATGAGC 2728
1111 1111 111 111 111 111 111 1111
3040 GAGGCTGACAAGGCGCAGAGGAGGTCAAGCTTTCCGTACTGGGACATCAATGAAG 3099
2729 AGCAGCCCTTACAGACCTGCTGTTGAGTTACTGATCAACCAAGAAAGCGGAAACCC 2788
1111 111 111 111 111 111 111 1111
3100 GCCATCGGACTCTTCATCTCTCTCAGCATCTTCTTTCATGTGTAAACATGTCTC 3159
2789 ACTGTGAC-----CGAGGGAAGAGACCTGGGTGAGTGAAGACATGAAGACATCT 2842
1111 1111 111 111 111 111 111 1111
3160 GCGCTGGCTTCCAACTATTTGGCTCAGCCTGTGACTGATGATGCCCTCGTCAACGGGACT 3219
2843 CATATGAGTACTATGCGAGCATACGCCCTCTCCATGGCAGTCACTGATCTGTA 2902

111 111 111 111 111 111 111 111
3220 CAGGAGACACACAAAGTCCGGCTGAGCGCTATGAGACCCCTGGCATTTTCACAAAGGATC 3279
2903 GCCATTGAG-----AGTGTCTTTGTCAAGGCGACGCTCGAGCTTCCTCCGGCTG 2956
111 111 111 111 111 111 111 1111
3280 GCCGTTTGGCTACTCCATATGGCCGTGTCATCGGGGGGATCTTGCTCCGCTGTCTC 3339
2957 CATGACAGCTTTTCCGAAGATTCCTGAAAGCCCTATGAATTTTTCACAGACCCCC 3016
1111 111 111 111 111 111 111 1111
3340 CAGGTGACCTCTGACAGCATCTGCGGTACCATGAGCTTCTTGTAGCGGACCC 3399
3017 ACAGGAGATTTCTCAACAGTTTTCCAAAGACATGATGAAATTGAGTGGCGCTGCG 3076
1111 111 111 111 111 111 111 1111
3400 AGTGGGAACCTGTGACACCGCTCTTCAGAGAGTGGACAGTGGACATCCATCCCG 3459
3077 TTCAGAGCCGAGATGTTCAATCCAGAACTTATCTGTGTTCTTCTGTGGAAATGATC 3136
1111 111 111 111 111 111 111 1111
3460 GAGTCATCAAGATGTCTCAGTGGCTCCGTTCACAGTCAATGAGTCCGATCTTATC 3519
3137 GCAGAGTCTTCCCGTGGTTCTTGTGGCAGTGGGGCCCTGTGATCTCTTTTCACTC 3196
111 111 111 111 111 111 111 1111
3520 CTGTGGCCAGCGCCATGCGCCCATCATCATCCGCCCTTGGCTCAATCTTCTTC 3579
3197 CTGACATTTCTCCAGGGTCCCTGATTCGGAGCTGAACGCTTGGACAAATATCAAGCAG 3256
111 111 111 111 111 111 111 1111
3580 GTCCAGAGTCTTACGTGGCTTCTCCCGGACGTGAACGCGCTGAGTGGTCAAGCCGC 3639
3257 TCACCTTCTCTCCCATCATCACTGACATCAAGGAGCTTGGCCATCCAGCGCTAC 3316
111 111 111 111 111 111 111 1111
3640 TCCCGGTCTATTTCCATTTCAACAGACATCTGTGGGGGTCAAGCTTCTGAGCCTTC 3699
3317 AATTAAGGACAGAGTCTGACAGATACAGAGCTGCTGATGACAAAGCTCT 3376
111 111 111 111 111 111 111 1111
3700 GAGGAGACAGACGCTTCATCCACAGAGTGAACCTGAAGGTGACAGAACAGAGGC 3759
3377 TTTTGTGTTTACGATGTCGATGGGTGGCTGTGCTGCGCTGACCTCATCAACATC 3436
111 111 111 111 111 111 111 1111
3760 TATTTACCCAGATGATGAGCAAGGTGCTGCGCTGAGTGGTGGCAAC 3819
3437 GCCCTCATCACACACAGGGGCTGATGATGCTTCTTATCACGGGAGATTCGCCACGC 3496
111 111 111 111 111 111 111 1111
3820 TCCATCTGTTCTGTTGCTCCCTGTTGGGTGATCTCCAGGCACAGCTCAGTGTGCG 3879
3497 TATGGGCTCTGCCATCTCTTATGCTGTGCAAGTAAACGGGGCTGTTCAGTTTACGTC 3556
111 111 111 111 111 111 111 1111
3880 TTGCTGGCTCTCACTGCTTACATTCATTCAGGTACACAGTACTTGAACCTGCTGCT 3939
3557 AGACTGCATCTGAGACAGAACTGATTCACCTGCGTGGAGAGATCAATCAAT 3616
111 111 111 111 111 111 111 1111
3940 CGGATGTCATCTGAATGGAACCAACATCGTGGCGTGAAGGCTCAAGGATTTCA 3999
3617 AAGACTCTCTTGGAGACCTGCGCAGAAATTAAGAACAGCTCTCCCTGACTG 3676
1111 111 111 111 111 111 111 1111
4000 GAGACTGAGAA---GAGAGCGCCCTGGCAAAATCCAGGAAACAGCTCCGACAGCTGG 4056
3677 CCCAGGAGGAGAGGTGACTTTTGAACGACAGATGAGGATGAGGAAAGAAACCTCCCT 3736
1111 111 111 111 111 111 111 1111
4057 CCCAGGTGGCGGAGTGAATTCGGAACTACTGCTCGCTACCGAAGAGACCTGGAC 4116
3737 CTTCCTTAAAGAAATATCTTCAAGATCAAACTTAAGAGAAATTTGGATGGG 3796
1111 111 111 111 111 111 111 1111
4117 TTCTGTTCTCAGACATCAATGATCAAGATCAATGAGGAGAAAGTCCGATCGGGG 4176
3797 CGGACAGATCAAGGGAATCTCGCTGGGGAATGCGCTTCTCCGTGTGGTGAATATCT 3856
1111 111 111 111 111 111 111 1111
4177 CGGAGCGAGCTGGGAGTGCCTGAGCCCTGGGCTTATTTTCGATCAACAGATCTGCG 4236
3857 GGAGGCTGATCAAGATATGATGAGTGAATCAAGATGAGTATTTGCTCCGACCTCGA 3916
111 111 111 111 111 111 111 1111
4237 GAAAGGAGATATCATGATGATCAACATGCGCAATGCGCTCGACAGACCTCGC 4296
3917 AGCAACTCTATCATCTTCAAGAGCGGCTGCTTCAAGTGGCACTGTCAATCAAT 3976
111 111 111 111 111 111 111 1111

Db 4297 TTGACATCACCATCATCCCCCAGAGACCCCTGTTTGTTCGGGTTCCTCCGATGAC 4356
Qy 3977 TTGACCCCTTCAACCACTACCTGAGACAGATTGGATGCCCTGGAGAGACAC 4036
Db 4357 CTGACCCATTCAGCCGCTGATGAGAGAGTGTGAGAGCTGAGAGCTGGAGCTGCCAC 4416
Qy 4037 ATGAAAGATGTATGTCTGACGCTACCTCTGAAACTTGAATCTGAAAGTGAAGTGG 4096
Db 4417 CTGAAAGACTTCGTGACCCCTTCGACAAAGTAAAGTAAATGAGAGAGGGGG 4476
Qy 4097 GATACCTTCAGTGGGGGAGAGCAGCTCTTGATAGAGCCCTGCTCCGCCAC 4156
Db 4477 GAGAACCTCAGTGTGCGGAGCCGCTGCTGTGCTAGCCCGGCTGCTGAGAGAG 4536
Qy 4157 TGTAAATCTGATTTTGAATGAGACCACTGCTGATGACACAGAGAGACTTTTG 4216
Db 4537 ACGAAGATCTTGTGTGATGAGAGCCAGGACCGCTGAGACCTGGAAGAGAGACTC 4596
Qy 4217 ATTCAGAGACCATCCGAGAGACTTTGACAGACTGTACATGCTGACCATTTGCCATCC 4276
Db 4597 ATCCAGTCCACATCCGAGACAGATTGAGAGACTGACACCGCTCCACCATCCGACCGG 4656
Qy 4277 CTGACACGCTTCAGGCTCCGATAGATATGCTGCTGGGCCAGGAGAGAGTGGAG 4336
Db 4657 CTGACACCATCATGAGCTACACAGAGGTGATGCTTTGGACAAAGAGAGATCCAGAG 4716
Qy 4337 TTTGACACCCCATCGTCTCTCTG 4360
Db 4717 TACGGCGCCCATCGAGACTCTG 4740

RESULT 7
US-08-463-179A-3
Sequence 3, Application US/08463179A
Patent No. 6001563
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,179A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConit, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: P01-002CP8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 5011 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 196..4788
US-08-463-179A-3

Query Match 10.5%; Score 511.2; DB 3; Length 5011;
Best Local Similarity 52.5%; Pred. No. 1,4e-122;
Matches 1367; Conservative 0; Mismatches 1178; Indels 59; Gaps 9;

Qy 1811 CACATCCACCTGGGCGCCACCTGCTTACAGAGACACTGACACATGATCTGAGATC 1870
Db 2142 CACATTCACCTGGGCGCA--GGAGCGACCTCCACACTGAATGSCATCACCCTTCATC 2139
Qy 1871 CAGAGGGTAAACTGTTGAATCTGGGCGAGTGGGAAAGTGAAGAAACCTCTCATTT 1930
Db 2200 CCGAAGGTGCTTGTGTGGCGCGTGGTGGCCAGGTGGGCTCGGAAAGTGCCTCCTC 2259
Qy 1931 TCACCATTTTGGCCAGAGTACCTCTTGAAGGCGACATGTCATTCAGTGAACCTTC 1990
Db 2260 TCACCCCTTGGCTGAGATGACAAAGTGAAGGGGACGCTGCTATCAAGGCTCCGTG 2319
Qy 1991 GCTATGTGGCGCCAGAGCGCTGGATCTCTCATGCTGAGAGACATCTCTTT 2050
Db 2320 GCTATGTGGCGCCAGAGCGCTGGATCTCTCATGCTGAGAGACATCTCTTT 2050
Qy 2051 GGAAGGAATATGATGAAGAAAGATACAACTGTGCTGAAGACCTGCTGAGGCT 2110
Db 2380 GGATGTGAGCTGAGAGAACCATATTAAGGCTGATACAGGCTGCTGCTCTCCCA 2439
Qy 2111 GACCTGGCCATCTTCCACAGAGAGCCTGAGAGATGAGAGAGAGAGAGAGAGAG 2170
Db 2440 GACCTGGAATCTGCGCCAGAGGAGTGGACAGAGATGAGAGAGAGAGAGAGAGAG 2439
Qy 2171 AGCGTGGGCGAGCGCCAGAGAGATGAGCTTGGCGGCTGATGAGAGAGAGATC 2230
Db 2500 TCTGGGGAG 2559
Qy 2231 TACATCTGAG 2290
Db 2560 TACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2619
Qy 2291 AGTGTAT-----CGGAAGATCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 2344
Db 2620 AATGTATGAGCGCCAG 2679
Qy 2345 TTACAGTACCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2404
Db 2680 ATGAGTACCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2739
Qy 2405 AGAGCAGCCATGAG 2464
Db 2740 ATGGGCTCTTACAG 2799
Qy 2465 CTGTGCTGAG 2521
Db 2800 TATCCAG 2859
Qy 2522 TCACAG 2573
Db 2860 CCAG 2919
Qy 2574 AAAAAAG 2615
Db 2920 CAACTGAG 2979
Qy 2616 AAGAGAAAGGAG 2675


```

3497 TATGGGCTGTCGTCATCTTATGTCAGTTAACGGGCTGTTCAGTTTAAAGGTC 3556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3880 TTGGTGGGCTCTCAGTGTCTTACTTACATTCAGGTACACAGCTACTTGAAGTGGTT 3939
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3557 AGACTGCATCTGAGACAGAACCTGATTCACCTCGGTGAGAGAGATCAATCAATCAAT 3616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3940 CGGATGCATCTGAATGAATCAACAACTCGTGGCCGTGAGAGAGGTCAAGAGATTCA 3999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3617 AAGACTCTGTCTTGGAGACCTCGCAGAAATTAGAACAAAGCTCCCTCCCTGATGG 3676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4000 GAGACTGAGAA--GGAGGCGCCCTGGCAAAATCCAGAGACAGCTCCGCCAGAGCTGG 4056
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3677 CCCGAGGAGGAGAGAGTACCTTTGAAACGAGAGATGAGTACCGAGAAACCTCCCT 3736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4057 CCCGAGGAGGAGGAGTGAATTCGGGAATCTACTGCTTCGCTACCGAGAGAGCTGGAC 4116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3737 CTGCTCTTAAAGAAATCTTCAAGATCAAACTTAAAGAGAAATTTGGATTTGGGG 3796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4117 TTGCTTTCAGGACATCAATGTCAAGATCAATGAGGAGAAAGTGGCATCGTGGGG 4176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3797 CGGACAGATCAAGGAATCTCTGCGGAGATGCGCCCTCTTCCGTGTGATGATATCT 3856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4177 CGGAGGGGAGCTGGAGAGTCTGCTGACCCCTGAGCTTATTCGATCAACGAGCTGCC 4236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3857 GGAGGCTGCAATGATGATGAGATGAGAAATCAAGATGATGAGCTTCCGAGCTCCGA 3916
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4237 GAGAGAGAGATCATGATGAGATGAGATCAACATGAGATGAGATGAGATGAGATGAG 4296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3917 AGCAACTCTCTATCTATCTCTCAAGAGGCGGTCTGTCAGTGGAGATGTCAGATCAAT 3976
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4297 TTCAAGATTCACATCATGATCCCAAGAACCTGTTTGTGTTGCGGGTTCCCTCGAATGAA 4356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3977 TTGAGCCCTTCAACAGTACTGTAAGACAGATTTGGGATCCCTGAGAGACACAC 4036
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4357 CTGGAGCCATTCAGCAGTCTGATGAGAAAGTCTGAGCTCCCTGAGAGTGGCCAC 4416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4037 ATGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4417 CTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4097 GATTAATCTCTAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4477 GAGAACCTCAGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4157 TGTAAATCTGATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4537 AGGAGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4217 ATTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4597 ATTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4277 CTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4657 CTCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4716
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4337 TTTGACACCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4717 TACGGGCGCCCATGAGAGCTCTG 4740
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

STREET: 60 STATE STREET, SUITE 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/141,893
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923; 08/029,340
FILING DATE: 27-OCT-1992; 8-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: P01-002
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5149
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5011 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 196..4788
US-08-141-893-1
Query Match 10.4%; Score 506.4; DB:1; Length 5011;
Best Local Similarity 52.4%; Pred. No. 2.4e-121;
Matches 1364; Conservative 0; Mismatches 1181; Indels 59; Gaps 9;
1811 CACATCCACTGGGCGACCTGCGCTTACAGAGACACTGCACAGATCGATGAGATC 1870
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2142 CACATCCACTGGGCGCA--GAGAGACCTCCACACCTGAATGGATCACTTCCTCATC 2199
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1871 CAGAGGCTAAACTGGTGAATCTGCGGAGTGTGGAGAGTGGAAAACTCTCTCAT 1930
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2200 CCCGAGGAGTGTGTTGGTGGCGCTGAGGCGCAGTGGGCTGCGGAAAGTTGCCGCTC 2259
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1931 TCAGCATTTTGGCGCAAGTGAAGCTTCTAGAGGAGAGATGAGATGCAATCGTGAACCTTC 1990
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2260 TCAGCCTCTTGGCTGAGAGCAAAAGTGAAGGAGGACAGTGGCTATCAAGGAGCTCGTG 2319
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1991 GCTTATGTTGGCCAGAGAGCTGAGATCTCTCAATGCTACTGTGAGAGACAATCTGTTT 2050
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2320 GCTTATGTTGGCCAGAGAGCTGAGATGAGATGATGATGATGATGATGATGATGATGAT 2379
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2051 GGGAGGAATATGATGAGAGAAAGATCAACTGTGCTGAGACAGCTGCTGAGGCT 2110
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2380 GGATGTCAGCTGAGAGAACATATTACAGGTCCGATGATCAAGGCTGCTGCTCCCA 2439
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2111 GACCTGGCATCTTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2170
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2440 GACCTGGAATCTTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2499
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2171 AGCGTGGCAGAGCGCAGAGATCAAGCTTGGCCGCGCTTGTATGATGAGAGAGATC 2230
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2500 TCTGGGGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2559
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2231 TACATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2290
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2560 TACCTTTCAGATGATCCCTCTCAGAGAGTGAATGCCATGATGAGAGAGAGAGAGAGAT 2619
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```


Db 3340 CACGTGACCTGTCACAGATCTCGGTCACCCATGAGCTCTTTGAGCGGACCC 3399
OY 3017 ACAGGAGGATTTTCACAGGTTTTCAGAGACATGATGATGAGTGCCTGGCTGGC 3076
Db 3400 AGTGGAGCTGTGTAACCGTTCTCCAGAGGCTGGACACAGTGAGTCCATGATCCG 3459
OY 3077 TTCAGGCGGAGATGTTTCATCCAGAGGATTAATCTGTTCTCTGTGAGGATGATC 3136
Db 3460 GAGGTATCAAGATGTTGATGAGGCTCCGTTTCAACGTCATGTTGCTGATGATC 3519
OY 3137 GCAGAGATCTCCCGTGTCTTGTGAGTGGGCCCCCTTGATCTCTTTTCACT 3186
Db 3520 CTGTGCGACGCGCATGCGCGCATCATATCCGCCCTTGCCCTCATCTACTTCTC 3579
OY 3197 CTGCAATGTCCTCCAGGTCCTGATTTGGGAGTGAAGCTGCAATATACAGAG 3256
Db 3580 GTCCAGAGTTTCTAGTGCTCTCCCGGACGCTGAAGCGCTCGAGTGGGTACGCC 3639
OY 3457 TCACCTTCTCTCCACATCCATCCAGATACAGGCGCTTGCCACATCCAGCGCTAC 3316
Db 3640 TCCCGGCTATTCCTCCATTCACGAGACCTTGCTGGGGGTACAGCTCATTCGAGCC 3699
OY 3317 AATTAAGGCGAGGATTTTCACAGATACAGAGCTGCTGATGACACCAAGCTCT 3376
Db 3700 GAGAGAGGAGCGCTTCATCCACAGATGACCTGAAGGTGACGAGAACCAAGGCC 3759
OY 3377 TTTTCTTTTCTTACGTCGATGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 3436
Db 3760 TATTAACCCAGCATCTGAGGCAACAGGTGGTGGCTGGCTGGCTGGCTGGCTGG 3819
OY 3437 GCGCTCATCACACAGCGGGCTGATGATGCTTTATGACGCGGACAGATCCCGAC 3486
Db 3820 TGCATGCTTCTGTTGCTGCGCTTTGCGGTGATCTCCAGGACAGCGCTGATGCT 3879
OY 3497 TATGCGGCTCTCGGCTCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3556
Db 3880 TTGGTGGCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3939
OY 3557 AGACTGGCTCTGAGACAGATGCTGATTCACCTGCTGAGGATGATGATGATGAT 3616
Db 3940 CGATGTCATCTGAATGAAACCAATGCTGGCCGCTGAGAGGCTGCAAGATATTCA 3959
OY 3617 AAGCTCTGCTTGAAGACCTGCGCAATTAAGAACAGCTCCCTCCCTGATGCTG 3676
Db 4000 GAGACTGAGAA--GAGGCGCGCTGCGCAATCCAGAGACCTCCGCGAGCAGCTGG 4056
OY 3677 CCGCAGGAGGAGAGGAGCTTGAAGCGAGATGAGTACCGAGAAACCTCCCT 3736
Db 4057 CCGCAGGTGGCGGAGTGAATCCGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 4116
OY 3737 CTGCTCTTAAAGAAATATCTTCACGATCAAACTTAAGAGAAATGAGATGAGG 3796
Db 4117 TTGCTTTCAGGCAATCAATGTCACGATCAATGAGGAGAAAGTGGGATGTTGG 4176
OY 3797 CGGACAGATGAGGAGGCTGCTGCTGAGGATGCTCTTCCCTGCTGAGTATCT 3856
Db 4177 CGGAGGAGGAGTGGAGTCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4236
OY 3857 GGAGGCTGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 3916
Db 4237 GAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4296
OY 3917 AGCAATCTCTATCATCTCTCAAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTG 3976
Db 4297 TTTCAAGATCAACATCAATCCCGAGACCGCTTTTGTGTTGCGGTCCCTCGATGAC 4356
OY 3977 TTGACACCTTCAACGATCACTGAAGACCAAGATTTGGAGTCCCTGAGAGACAC 4036
Db 4357 CTGAGCCCATTCACGCTGCTGATGAGAACTGATGAGCTGAGCTGAGCTGAGCTG 4416
OY 4037 ATGAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4096

Db 4417 CTGAGGACTCTGCTGACGCCCTTCTGACAGCTAGACATGATGTGAGAGCGGG 4476
OY 4097 GATACCTTCAGTGGGGGAGAGCGACCTTGTGATAGTAGAGCCCTGCTCCGAC 4156
Db 4477 GAGAACTCACTGTGCGGAGCGGCGAGCTTGTGCTAGCGGGCCCTGCTAGAGAG 4536
OY 4157 TGTAAATCTGATTTTGAAGAACCAAGCTGCTGATGAGACAGAGAGACTTATG 4216
Db 4537 ACGAAGATCTCTGTTGATGAGAGCGAGCGGCGGCTGAGCTGAGAAAGAGAGCTC 4596
OY 4217 ATTCAGAGACCTCGAGAAAGATTTGACAGCTGATGATGATGATGATGATGATG 4276
Db 4597 ATTCAGAGACCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4556
OY 4277 CTGACAGGTTTCTAGGCTCGATGATGATGATGATGATGATGATGATGATG 4336
Db 4657 CTCAACACATCATGAGCTACAGAGGATGATGATGATGATGATGATGATGATG 4716
OY 4337 TTGACACCCCATGCGTCTCTG 4360
Db 4717 TACGCGCGCCCATGAGCTCTCTG 4740

RESULT 11
US-08-462-109A-1
; Sequence 1, Application US/08462109A
; Patent No. 5882875
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,109A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Decont, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: P01-002CP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:

QY	2405	AGAGGACCCCATGAGGAACTGATGAAATTTAAATGGGATATGCTACCATTTTAAATAC	2464
Db	2740	ATGGGCTCTCTACAGAGAGCTGCTGCTCGAGACGGCGCTTGCTGAGTTCCTGCTAC	2799
QY	2465	CTGTGCTGGGAGAGACACCCGCAATT---GAGATCAATTCAAAAGGAAACAGTGT	2521
Db	2800	TATGCCAGCAGACAGGCGAGGAGCGAGTGTGAGAGGAACGGGGGTCACGGGCGTCA	2859
QY	2522	TCACAGAGAAGTTCACAGACAAG-----GTCTTAAACAGATATCAATAAAGAG	2573
Db	2860	CCAGGAGGAGGAGACCAAGCAAAATGGAATGGCATGCTGTTACGAGACAGTGCAG	2919
QY	2574	AAAAAGAGTAAAGCCAGAGGAAGGGCAGCTGTGAGCTG-----	2615
Db	2920	CAATCGAGAGACAGCTCAGCAGCTCTCTCTCTATATGAGGGAGCATCAGCAGCAC	2979
QY	2616	AAGAGAAAGGGCAGGGTTCAGTCCCTGGCTAGTATATGTGTCTACATCCAGGCTGCT	2675
Db	2980	AACAGCACCCGAGACTGCAAGAAAGCTTGAGGCCAGAGAGAGAGAACTTGAAGCTG	3039
QY	2676	GAGG-----CCCTTGGCATTCGATTATATGAGCCCTTTTCATGCTGAATGAGC	2728
Db	3040	GAGGCTACAAAGGCGCAGACAGGAGGAGTCAAGCTTCCGTACTAGGAGTACATGA	3099
QY	2729	AGCAGCGCTTCACAGCAGCTGTGTGTGAGTACTGATCAAGCAAGAGGCGGACAC	2788
Db	3100	GCCATCCGAGCTTCATCTCTCCACAGATCTTCCTTATGTTGAAACCATGTCTC	3159
QY	2789	ACTGTGACT-----GGAGGAAACGAGACCTGGGTGAGTGAAGATGAAGCAATCT	2842
Db	3160	GCGCTGCTTCACATATTTGTGCTCAGCCTCTGTGACTGATGACCCATGCTCAAGGG	3219
QY	2843	CATATGAGTACTATGCGACAGCATCTAGGCCCTCTCCATGAGCAGTATGCTGTA	2902
Db	3220	CAGAGACACAGAAAGTCCGGCTGAGCGTCTATGAGAGCCCTGGGATTTCAAGAG	3279
QY	2903	GCCATTGAG-----AGTTGCTTTGTCAAGGGCAGCGCTCGAGCTTCTCCGGCTG	2956
Db	3280	GCCGTGTTTGGCTACTCCAGAGGCCGTGTCATCCGGGGGATTTGGCTCCGCTGT	3339
QY	2957	CATGACAGCTTTCCGAGAGATCTTCGAAGCCCTATGAATTTTTCACAGCGACCC	3016
Db	3340	CAGCTGACCTGCTGCACAGCATCTCGCGGTACACCATAGGCTCTTTGAGCGGAC	3399
QY	3017	ACAAGGAGATTTCTCACAGAGTTTCCAAAGACATGATGAAGTTGACGTGGCGCTG	3076
Db	3400	AGTGGGAAGCTGGTGAACCGCTCTCCAAAGAGAGCTGGAACAGTGGACTCATTGAT	3459
QY	3077	TTCCAGGCCAGATGTTCAATCCAGAAAGTTATCTGTGTTCTTGTGTGGAAATG	3136
Db	3460	GAGGTCAATCAAGATGTTCAATGAGGCTCCCTGTTCAAGTATGTGCTCGATCT	3519
QY	3137	GCAAGAGTCTTCCCGTGTCTCTGTGGCAGTGGGGCCCTGTATCTCTTTTCACT	3196
Db	3520	CTGCTGGCACGGCCATCGCCGCATCAATCAGCCGCCCTTGGCTCTCATCTTCTTC	3579
QY	3197	CTGCAGATTTCTCCAGGGTCTGATTTCCGAGGAGCAACGTCCTGGACAATATCA	3256
Db	3580	GTCACAGAGTTTACAGTGGTCTCTCCCGGACGCTGAAGCGCTCGAGTGGCTAG	3639
QY	3257	TCACCTTTCTCTCCACATCAGCTCCAGATACAGGGCGTTTCCACCATCCAGCTAC	3316
Db	3640	TCCCGGCTATTTCTCCATTTTCAACGAGACCTTCTGGGGGTGAGGTATTCAGAC	3699
QY	3317	AATTAAGGGAGGATTTCTGACAGATACAGAGAGCTGCTGGATGAACAACCAACT	3376
Db	3700	GAGAGCAGGAGCGCTTATCCACACAGAGTACCTGAAGGTGACAGAACCAAGAGCC	3759
QY	3377	TTTTTTTGTTAAGGTGAGTGCAGTGGAGTGGCTGTGCTGGCTGAGACTCATCA	3436
Db	3760	TATTAACCAAGCATCTGAGGCAACAGTGGTGTGCGCTGTGAAGTGTGGGCAAC	3819
QY	3437	GCCCTCATCACACAGGGGGTGTATGATGTTTATGACAGGGGCAAGTTCCCCAGCC	3496

Db	3820	TGCATCTCTCTGTTTGGCTGGCCGTTTGGGGTGATGTCACAGGCACAGCCTCAGTCTGGC	387.9
Oy	3497	TATCGGGGTCTCGGCATCTCTTAATGCTGTCAGTTAAGGGGCTGTCAGTTTACGGTC	355.6
Db	3880	TTGGTGGGGCCTCAGTGTCTTACTACTATGACAGTGCACCACTACTTGAACTGGCTGGTT	393.9
Oy	3557	AGACTGGCATCTGAGCAGAAAGCTGCATTCACCTGGGTGGAGAGAGTAATTAATACACTNT	361.6
Db	3940	CGGATGTCATCTGAAATAGGAACCAACATCTGGCCGTGTGAAGAGGCTCAAGAGATTTC	399.9
Oy	3617	AAGACTCTGTCTCTGGAAGACCTGCACAGATTAAACAAAGGCTCCCTCCCTACTACG	367.6
Db	4000	GAGACTGAGAA - - - GAGAGGGCCCTGGCAAAATCCAGNAGACACAGTCCGGCCAGCAGCTGG	405.6
Oy	3677	CCCCAGGAGGAGAGTGACCTTTGGAACGCAGAGATGAGTACCGAGAAAACTCCCT	373.8
Db	4057	CCCCAGGTGGGCCGAGTGGAAATTCGGGAACACTACGCTCGCTACCGAGAGACCTGGAC	411.8
Oy	3737	CTCTGCTTAAAGAAAGTATCTCTCAGATCAAAACCTAAAGAAAGATTGGGCTTTGGGG	379.8
Db	4117	TTTCGTTCTCAGGGCATCAATATGTCAGATCAATAGGGGGAGAAAGGTGCGCATCTGGGG	417.6
Oy	3797	CGCAGAGATCAGGAGAGTCTCGTGGGGATGGCCCTTCCGCTGTGGTGGAGTTATCT	385.6
Db	4177	CGGACGGGAGCTGGGAAGTGTGCTCCAGACCCTGGGCTATTTCGGATCAACAGATCTGCC	423.8
Oy	3857	GGAGGCTGCATCAAGATTGATGAGATGAGATCAGTATATTTGGCTTGCAGACTCCGA	391.6
Db	4237	GAAAGAGAGATCAATCATGATGGCATCAACATCGCCAAATGGGCTGCAGCACCTCCGC	429.8
Oy	3917	AGCAACCTCTCATTCATTCCTCAAGAGCCGGTCTGTAGTGGGCACTCTCATGATCAAT	397.6
Db	4287	TTCAAGATCAACATCATCACCACAGACCCGTGTTTGTTCGGGTTCCCTCGGAATGAC	435.6
Oy	3977	TTTGACCCCTTCAACACAGTACACTGAAGACAGATTGGGATGCCGTGAGAGACACAC	403.8
Db	4357	CTGGACCCATTCACCCAGTACTCGGATGAAGAAGTGTGGACCTCCCTGGAGCTGGCCAC	441.8
Oy	4037	ATGAAAGATGTATGTCTCAGCTACCTCTGAAACTTGAAATCTGAAATGTATGAGAAATGG	409.8
Db	4417	CTGAAGGACTTCGTGTCAGCCCTTCTCGAACACTGACCAATGTCACAAAGGCGGG	447.6
Oy	4097	GATTAATCTTCAGTGGGGGAAAGCGGACACTTTTGTGATAGCTAGAGCCCTGCTCGCCAC	415.6
Db	4477	GAGAACTCAGTGTGGGGCAGCCGACAGCTGTGTGCTTAGCCCGGGCCCTGCTGAGGAAG	453.8
Oy	4157	TGTAAAGATTCTGATTTTATGATGAAGCCACACAGCTGCCATGAGACACAGACAGACTTATTG	421.8
Db	4537	ACGAAATCTCTGTGTGTATGATGAGCCACGCGCAGCCGTGAGACTGTGAAGACGACACTC	459.8
Oy	4217	ATTCAGAGACATCCGAGAAAGCATTTGGCATGTATACCATGCTGTGACCTTCCCATCGC	427.6
Db	4597	ATTCAGTCCACATCCGGACACAGTTGTGAGAGATGACGTCCTCTACCATGCGCCACCGG	465.6
Oy	4277	CTGCACACAGGTTTAAAGCTCCGATAGGATTATGTGCTGGCCACGAGACAGGTGTGGAG	433.6
Db	4657	CTCAACACACATCATGAGCTACACAAAGGATGATGCTTGTGACAAAGAGAAATTCAGAGAG	471.8
Oy	4337	TTTGACACCCCATGGTCTCTTGG 436.0	
Db	4717	TACGGCGCCCATGGAGCTCTCG 474.0	

RESULT 13 179A-1
US-08-463-179A-1
: Sequence 1, Application US/08463179A
: Patent No. 6001563
: GENERAL INFORMATION:
: APPLICANT: COLE, Susan P.C.
: APPLICANT: Deley, Roger G.
: TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
: NUMBER OF SEQUENCES: 6


```

Db 3580 GTCCAGAGGTTCTAGTGGCTTCTCCGCGGACGCTGAAGCGCTGAGTGGTACGCCG 3639
QY 3257 TCACCTTCTCTCCACATACGTCAGCATACAGGCGCTTCCACCATCCAGCGCTAC 3316
Db 3640 TCCCGCGTCTATCCATTCATCAAGAGACCTTGGCGGGGTACGCGTCAATCCAGCTTTC 3699
QY 3317 AATAAGGCGAGGATTCCTGACAGATACAGAGCGCTGCGATGACAACTGCTCT 3376
Db 3700 GAGGAGCAGAGGCGCTTATCCACAGAGTACCTGAAGGTGACGAGAACGAGAGGCC 3759
QY 3377 TTTTGTGTTTACGTCGTCATGCGGTCGGTGGCTGCGGCTGACCTCATACATC 3436
Db 3760 TATATCCACAGATCGTGGCCACAGGTGGCTGGCGCTGCGGCTGAGTGTGGGCAAC 3819
QY 3437 GCCCTCATCCACCAGCGGGGCTGATGATCGTTTATGACAGGGGAGATTCGCCAGCC 3496
Db 3820 TGCATCGTTCTGTTTGTGCTGCGCTGTTGCGGTGATCTCCAGGACAGCGCTCAGTCTGGC 3879
QY 3437 TATGGGGGTCTGCGCATCTCTATGCTGCTCCAGTTACGGGGCTTTCAGTTACGGTC 3556
Db 3880 TTGCTGGGCTCTCTCATGCTCTTACATTCAGTCCAGTCCAGCATCTGAATCGCTGGTT 3939
QY 3557 AGACTGGCATCTGAGACAGAAAGCTGATTCACCTCGGTGAGAGAGATCATACAT 3616
Db 3940 CGGATGTCATCTGAATGTAAGAAACCAACATCGTGCGCTGAGAGAGGCTCAAGAGATTTCA 3999
QY 3617 AAGACTGTCTCTGGAAGACCTGCCAATTAAGAACAAAGGCTCCCTCCCTGACTGG 3676
Db 4000 GAGACTGAGAA---GGAAGCGCGCTGGCAAAATCCAGAGACACGTCGCCCGCAGCACTGG 4056
QY 3677 CCCGAGGAGGAGAGTACCTTTAGAAACGAGAGATGAGTACCAGAAACCTCCCT 3736
Db 4057 CCCCAGGCGGCGGAGTGAATCCGGAATCTGCGCTGCGTACGAGAGAGACCTGGAC 4116
QY 3737 CTCGCTCTAAAGAAAGTATCTTCACGATCAAACTAAAGAGATGGGCTATGTTGGG 3796
Db 4117 TTCGTTCTCAGGACATCATATGTCTACGATCAATGGGGGAGAAAGTGGGCTGCTGGG 4176
QY 3797 CGGACAGATCAGGAGTCTCTGCTGGGGATGGCCCTCTTCGCTGCTGGTGAATATCT 3856
Db 4177 CGGAGGGAGGCTGGGAAATCGGCTGACCCCTGAGCCCTGATTTCCGATCAACGATCTGCC 4236
QY 3857 GGAGCTGCATCAGATGATGATGAGTGAATCATGATATGGCTTGGCCACTCCGA 3916
Db 4237 GAAGGAGATCATCATGATGATGATCAACATCGCAAGATGGCGCTGCACACCTCCGC 4296
QY 3917 AGCAAACTCTATCATCTCTCAAGAGCGGCTGCTCAGTGGGACATGTCAGATCAAT 3976
Db 4297 TTCAGATCATCATATCCCCAGAGACCTGTTTGTGTTTGGGTTCCCTCGAATGAAC 4356
QY 3977 TTGACCCCTTCAACAGTACACTGAAGACAGATTTGGGATGCCCTGGAGAGACACAC 4036
Db 4357 CTGAGCCATTCAGCGAGTACTCGGATGAAGAAATCTGGAGCTCCCTGGAGTGGGCCAC 4416
QY 4037 ATGAAGATGATGATGCTGACGTAACCTCTGAACCTTGAATCTGAATGATGGAATGG 4096
Db 4417 CTGAAGGACTTGTGTCAGCGCTCTCTGACAGCTAGACATGAATGTGCAAAAGCGGG 4476
QY 4097 GATACCTTCTCAGTGGGGAACGGCAGCTCTGTGATCATAGAGGCCCTGTCCGCCAC 4156
Db 4477 GAGAACTCAGTGTGGGCGGACGCTGTGTGTGCTGAGCCCGGCGCTGTGAGGAAG 4536
QY 4157 TGTAAAGTCTGATTTTATGATGAAGCACAGCTGCATGAGACAGAGACAGATTTATG 4216
Db 4537 ACGAATCATCTGTGTGATGATGAGCGACAGCGGACCGTGAAGCGAGACGACCTC 4596
QY 4217 ATTCAAGACATCTCGAGAAAGATTTGACAGCTGTACATGCTGACCTTGGCCATGGC 4276
Db 4597 ATCCAGCTCACCATCCGAGACAGATTTGAGGAGTCTACACGCTCTCAACATCCGCCAG 4656
QY 4277 CTGACAGGCTCTGAGGCTCCGATAGATGATGATGCTGGCCAGGAGACAGTGGTGGAG 4336

```

```

Db 4657 CTCACACCATCATGAGTACACAAAGGATGATGTTGACAAAGAGAAATCCAGGAG 4716
QY 4337 TTGACACCCCATCGTCTGCTGTG 4360
Db 4717 TACGGCGCCCATCGACCTCTCTG 4740

RESULT 14
US-08-461-384B-1
; Sequence 1, Application US/08461384B
; Patent No. 6025473
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P. C.
; APPLICANT: Dealey, Roger G.
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,384B
; FILING DATE: 05-JUN-95
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Steed, Carol Mlenickl
; REGISTRATION NUMBER: 39,539
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 196..4788
; US-08-461-384B-1

Query Match 10.4%; Score 506.4; DB 3; Length 5011;
Best Local Similarity 52.4%; Pred.No. 2.4e-121;
Matches 1364; Conservative 0; Mismatches 1181; Indels 59; Gaps 9;
QY 1811 CACATCACTGCGGCGCTTACAGAGACACTGACAGCATGATCTGAGATC 1870
Db 2142 CACATTCACCTGGGCA--GAGCGACCCCTCCACACATGAATGGATCATCTTCACATC 2199
QY 1871 CAAGAGGTAACCTGGTGAATCTGCGCAGTGTGGAGAGTGAAGAAACCTCTCATTT 1930
Db 2200 CCCGAAGGTGCTTGTGGTCCCTGTGTGGGCGAGGTGGGCTGGGAAAGTTGTCCTGCTC 2259
QY 1931 TCAGCATTTTGGCGCAATGACGCTTTAGAGGGGAGCATTTGCAATCATGTGAACCTTC 1990

```


THIS PAGE BLANK (USPTO)

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

QY	3668	CTGTACGAGCCCGAGAGAGAGAGTGCCTTTTGAGACCGCAGAGATGAGTACCGAGAA	3727
Db	661	CTGTACGAGCCCGAGAGAGAGTGCCTTTTGAGACCGCAGAGATGAGTACCGAGAA	720
QY	3728	AACCTCCCTCTGCTCTTAAGAAAGTATCTTCACGATCAAACTTAAAGAGAAGTGGC	3787
Db	721	AAACCTTCCTCTGCTCTTAAGAAAGTATCTTCACGATCAAACTTAAAGAGAAGTGGC	780
QY	3788	ATTGTGGGCGGGAAGAGATCAGGGAAGTCCGCTGGGGATGAGCCCTCTCCGTGGTG	3847
Db	781	ATTGTGGGCGGGAAGAGATCAGGGAAGTCCGCTGGGGATGAGCCCTCTCCGTGGTG	840
QY	3848	GAGTTATCTGGAGGCGTCATCA	3869
Db	841	GTTATCTTGGAGGCTTGCATCA	862
RESULT	2		
LOCUS	AUI37090		
DEFINITION	AUI37090	761 bp	mRNA
ACCESSION	AUI37090		EST
VERSION	AUI37090.1		24-OCT-2000
KEYWORDS	sequence.		
SOURCE	Placental tissue from Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 761)		
AUTHORS	Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isigaki, T.		
TITLE	HRI human cDNA project		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Takao Isigaki Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3951 Fax: 81-438-52-3952 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.		
FEATURES			
source	Location/Qualifiers		
	1..761		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="PLACE1005753"		
	/clone_lib="PLACE1"		
	/tissue_type="placenta"		
	/note="Vector: pELISF13"		
BASE COUNT	200 a	185 c	204 g
ORIGIN	165 t	7 others	
Query Match	14.2%	Score 686.6;	DB 108; Length 761;
Best Local Similarity	96.4%	Pred. No. 8.8e-162;	
Matches 721; Conservative	0;	Mismatches 24;	Indels 3; Gaps 2;
QY	3650	AAGAAACAGGCTCCCTCCCTGACGAGGCCCGCAGAGAGAGAGTACCTTTGAGAAGCA	3709
Db	1	AAGAAACAGGCTCCCTCCCTGACGAGGCCCGCAGAGAGAGAGTACCTTTGAGAAGCA	60
QY	3710	GAGATGAGGTACGAGAAACCTCCCTCTGCTCTTAAGAAAGATTCCTTCACGATCAA	3769
Db	61	GAGATGAGGTACGAGAAACCTCCCTCTGCTCTTAAGAAAGATTCCTTCACGATCAA	120
QY	3770	CTTAAAGAAAGATTGGCATTTGTGGGGGAGACAGATAGGGAATCTCGTGGGGATG	3829

Db	121	CCTAAGAGGAGATTGGCATTTGTGGGGGGGACAGAGATCAGGGAAGTCTCTCGTGGGGATG	180
QY	3830	GCCTCTTCCTCTCTGTGTGAGATTATCTGTGAGGCTGTGATCAAGATTGATGAGTGAATC	3889
Db	181	GCCTCTTCCTCTCTGTGTGAGATTATCTGTGAGGCTGTGATCAAGATTGATGAGTGAATC	240
QY	3890	AGTATATTTGGCTTGGCCGACCTCCGAAAGCAAACTCTCTATCATTTCCCTCAAGAGCCGGTG	3949
Db	241	AGTATATTTGGCTTGGCCGACCTCCGAAAGCAAACTCTCTATCATTTCCCTCAAGAGCCGGTG	300
QY	3950	CTGTTCAGTGTGCTACTGTTCAGATCAAAATTTGGACCCCTTCAACACAGTACACGGAAGACCAG	4009
Db	301	CTGTTCAGTGTGCTACTGTTCAGATCAAAATTTGGACCCCTTCAACACAGTACACGGAAGACCAG	360
QY	4010	ATTTTGGGATGCTCCCTGTGAGAGGACACACATGAAAGATGTATTGCTCACTACTCTGAAA	4069
Db	361	ATTTTGGGATGCTCCCTGTGAGAGGACACACATGAAAGATGTATTGCTCACTACTCTGAAA	420
QY	4070	CTTGAATCTGAAAGTATGGAATGGAATGAGGGATTAATCTTCAGTGGGGGAACGGCAGCTCTTG	4129
Db	421	CTTGAATCTGAAAGTATGGAATGGAATGAGGGATTAATCTTCAGTGGGGGAACGGCAGCTCTTG	480
QY	4130	TGCATAGCTTAAGGCTCCCTGTGAGAGGACACATGAAAGATGTATTGCTCACTACTCTGAAA	4189
Db	481	TGCATAGCTTAAGGCTCCCTGTGAGAGGACACATGAAAGATGTATTGCTCACTACTCTGAAA	540
QY	4190	GCCATGACACACAGACAGACTTATTGATTCAGAGACATCCGAGAACGATTTTGACAGC	4249
Db	541	GCCATGACACACAGACAGACTTATTGATTCAGAGACATCCGAGAACGATTTTGACAGC	600
QY	4250	TGTACCATGCTGACCATTTGGCCATTCGCTGTGACACAGGTTTACGCTCCGATA-6GATTAT	4308
Db	601	TGTACCATGCTGACCATTTGGCCATTCGCTGTGACACAGGTTTACGCTCCGATA-6GATTAT	660
QY	4309	GGTCTGTGGCCGAGGAGACAGTGTGTGAGTTTGACACCCCATCG--TCTCTGTGTCAAC	4366
Db	661	GGGCTGTGGCCGAGGAGACAGTGTGTGAGTTTGACACCCCATCGGTTCTCTGNGTCCAAC	720
QY	4367	GACAGTTCCGAGTTCTTATGCCATGTTTG	4394
Db	721	GACAGTTCCGAGTTCTTATGCCATGTTTG	748
RESULT	3		
LOCUS	AL543173		
DEFINITION	AL543173 LTR NF1006_PL2 Homo sapiens cDNA clone CS0D1002YC13 5		
ACCESSION	AL543173		
VERSION	AL543173.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 889)		
REFERENCE	Li, W.-B., Gruber, C., Jesse, J., and Polayes, D.		
REFERENCE	Full-length cDNA libraries and normalization		
REFERENCE	Unpublished (2001)		
COMMENT	Contact: Genoscope		
COMMENT	Genoscope - Centre National de Sequencage		
COMMENT	Bp 191 91006 EVRY cedex - France		
COMMENT	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
FEATURES	location/Qualifiers		
SOURCE	1..889		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0D1002YC13"		
	/clone_1b="LTR_NF1006_PL2"		
	/tissue_type="placenta"		
	/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA		
	was primed with a NotI-oligo(dT) primer. Five prime end		
	enriched, double-stranded cDNA was digested with Not I and		

cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Little Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : filiang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 208 a 218 c 234 g 208 t 1 others

ORIGIN

Query Match 14.0%: Score 677.8; DB 106; Length 889;

Best Local Similarity 99.0%: Pred. No. 1.5e-159; Mismatches 7; Indels 0; Gaps 0;

Matches 682; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 28 GCGTGGCCGGTGTCTCTGAGACAGGGGCGCAGAAATCTGATGTGAACCTAACAGCTCG 87
 Db 1 GCGTGGCCGGTGTCTCTGAGACAGGGGCGCAGAAATCTGATGTGAACCTAACAGCTCG 60
 QY 88 TGAGCCCTGGAACTCCACTCAGAGAGATGAGATTCGACATAGAAAAAGATATAT 147
 Db 61 TGAGCCCTGGAACTCCACTCAGAGAGATGAGATTCGACATAGAAAAAGATATAT 120
 QY 148 CATCCCCAGTCTGGGATATAGAGTGTGAGAGAGAACACAGCACTTCTGGAGCGCAGAG 207
 Db 121 CATCCCCAGTCTGGGATATAGAGTGTGAGAGAGAACACAGCACTTCTGGAGCGCAGAG 180
 QY 208 AGACCGTGAAGATTCAGAGTTCAGAGAACTGCAGCCCTGGAAATGCCAAGATGCTTGA 267
 Db 181 AGACCGTGAAGATTCAGAGTTCAGAGAACTGCAGCCCTGGAAATGCCAAGATGCTTGA 240
 QY 268 AACAGACCCGAGCGGAGGCGCTCTCTTGTATGCTTCATGCACTTCAGCTAGAAAT 327
 Db 241 AACAGACCCGAGCGGAGGCGCTCTCTTGTATGCTTCATGCACTTCAGCTAGAAAT 300
 QY 328 CCTGGATGAGAGCATCCCAAGGAGAAATACCATGATGCTTGAAGCCAT 387
 Db 301 CCTGGATGAGAGCATCCCAAGGAGAAATACCATGATGCTTGAAGCCAT 360
 QY 388 CCGGACTACTTGCACAAACACAGACCCAGAGTGAACATCTGGGCTTTTTCCTGTATGAC 447
 Db 361 CCGGACTACTTGCACAAACACAGACCCAGAGTGAACATCTGGGCTTTTTCCTGTATGAC 420
 QY 448 TTTTTCGAGGCTTCTCTCTGAGGCGGCTGAGGCGCACAAGAGGGGAGCTCTCAATGGA 507
 Db 421 TTTTTCGAGGCTTCTCTCTGAGGCGGCTGAGGCGCACAAGAGGGGAGCTCTCAATGGA 480
 QY 508 AGAGCTGTGTCTCTCTGACAGCAGAGTCTTCTGACGTGAAGTGCAGAAAGCTAGAGAG 567
 Db 481 AGAGCTGTGTCTCTCTGACAGCAGAGTCTTCTGACGTGAAGTGCAGAAAGCTAGAGAG 540
 QY 568 ACTGTGCAAGAGAGCTGAATGAATGAGTGGGCGCAGAGCTGCTCCGTCGAAGGGTGT 627
 Db 541 ACTGTGCAAGAGAGCTGAATGAATGAGTGGGCGCAGAGCTGCTCCGTCGAAGGGTGT 600
 QY 628 GTGATCTTCTGCGGACAGAGCTCATCTGTCATCTGTCGTCCTGAGATGATACCACT 687
 Db 601 GTGATCTTCTGCGGACAGAGCTCATCTGTCATCTGTCGTCCTGAGATGATACCACT 660
 QY 688 GCGTGGCTTCAAGTGAAGCAGGCTTCATCG 716
 Db 661 GCGTGGCTTCAAGTGAAGCAGGCTTCATCG 689

RESULT 4

AL524917

LOCUS AL524917 960 bp mRNA EST 13-FEB-2001

DEFINITION AL524917 LTR_NFL003_NBC3 Homo sapiens cDNA clone CS00C009Yg21 5

ACCESSION AL524917

VERSION AL524917.1 GI:12788410

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 960)

AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. 960

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS00C009Yg21"

/clone_lib="LTR_NFL003_NBC3"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by life technologies. Contact : Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax : (1) 301 610

8371 Email : filiang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 234 a 231 c 270 g 224 t 1 others

ORIGIN

Query Match 13.7%: Score 662.8; DB 105; Length 960;

Best Local Similarity 98.8%: Pred. No. 9.2e-156; Mismatches 7; Indels 1; Gaps 1;

Matches 678; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 31 TGGCGGCTTGTCTCTGAGACAGGGGCGCAGAAATCTGATGTGAACCTAACAGCTGTGA 90
 Db 1 TGGCGGCTTGTCTCTGAGACAGGGGCGCAGAAATCTGATGTGAACCTAACAGCTGTGA 59
 QY 91 GCGGCGGAGTCTCCAGTGAAGATGAGATTCGATAGGAAAAAGATATATCAT 150
 Db 60 GCGGCGGAGTCTCCAGTGAAGATGAGATTCGATAGGAAAAAGATATATCAT 119
 QY 151 CCGGAGTCTGGGTATGAGAGTGTGAGGAGAGAACACAGCACTTCTGGAGCGCAGAGAGA 210
 Db 120 CCGGAGTCTGGGTATGAGAGTGTGAGGAGAGAACACAGCACTTCTGGAGCGCAGAGAGA 179
 QY 211 CCGTGAAGATTCAGAGTTCAGAGAACTGCAGCCGTTGGAATGCCAAGATGCTTGAANA 270
 Db 180 CCGTGAAGATTCAGAGTTCAGAGAACTGCAGCCGTTGGAATGCCAAGATGCTTGAANA 239
 QY 271 AGCAGCCGAGCGGAGGCGCTCTCTTGATGCCCTCAGTTCCTGAGCTCAGAAATCCT 330
 Db 240 AGCAGCCGAGCGGAGGCGCTCTCTTGATGCCCTCAGTTCCTGAGCTCAGAAATCCT 299
 QY 331 GGATGAGAGCATCCCAAGGAGGAAATACCATGATGCTTGAAGTGCCTGAAGCCCATCG 390
 Db 300 GGATGAGAGCATCCCAAGGAGGAAATACCATGATGCTTGAAGTGCCTGAAGCCCATCG 359
 QY 391 GACTACTTGCACAAACACAGACCCAGTGAACATGCTGGGCTTTTTCCTGTATGACTTT 450
 Db 360 GACTACTTGCACAAACACAGACCCAGTGAACATGCTGGGCTTTTTCCTGTATGACTTT 419
 QY 451 TTCGAGGCTTCTCTGAGGCGGCTGAGGCGCACAAGAGGGGAGGCTCAGTGAAGA 510
 Db 420 TTCGAGGCTTCTCTGAGGCGGCTGAGGCGCACAAGAGGGGAGGCTCAGTGAAGA 479
 QY 511 CGTGTGTCTCTGTCCAGACAGAGTCTTGTGACGTGAAGTGCAGAAAGCTAGAGAGACT 570
 Db 480 CGTGTGTCTCTGTCCAGACAGAGTCTTGTGACGTGAAGTGCAGAAAGCTAGAGAGACT 539

OY	571		GTGGCAAGAAAGAGCTGATGAAGAATTBGGCCACAGCGTCTTGCTCCTGCCAAGSGTTGTGTG	630
Dd	540		GTCGCACAAGAGAGCTGAAGTAAGTAGTGGGCGACAACGGCTGCTTCCCTCGAAGGGTTGTGTG	599
OY	631		GATCTTCGCCCCACCACAGGCTCATCTCTCCATCCGCTGTCGATGATCACAGCGCGCG	690
Dd	600		GATCTTCTTCTGCCCACACAGGCTCATCTCTTCCATCTGTCCTGATGATCACAGCGACTGGC	659
OY	691		TGGCTTCAGTGGACGACGCTTCATGG	716
Dd	660		TGCGTTCAAGTGGACCAAAATTTTCAGG	685
RESULT	5			
LOCUS	U66687		1206 bp mRNA EST 13-JUL-1999	
DEFINITION	HSH66687 Human Homo sapiens cDNA similar to ATP-binding cassette transporter, mRNA sequence.			
ACCESSION	U66687			
VERSION	U66687.1		GI:1906572	
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 1206)			
TITLE	Allikmets,R., Gerrard,B., Hutchinson,A. and Dean,M. Characterization of the human ABC superfamily: Isolation and mapping of 21 new genes using the expressed sequence tags database Hum. Mol. Genet. 5 (10), 1649-1655 (1996) 97049974			
JOURNAL MEDLINE	Contact: Dean, Michael National Cancer Institute, NCI-FCRDC Frederick, MD 21702, USA Email: dean@fcrlv1.ncifcrf.gov derived from EST: EST2771145, Location/Qualifiers			
FEATURES	source 1..1206 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="Human"			
BASE COUNT	273 a 284 c 275 g 362 t 12 others			
ORIGIN				
Query Match	13.0%; Score 628.6; DB 189; Length 1206; Best Local Similarity 93.6%; Pred. No. 4.1e-147;			
Matches	711; Conservative 0; Mismatches 37; Indels 12; Gaps 5			
OY	4054		TCAGCTACTCTGGAACCTTGATCTGTAAGTGAATGGGATAACTCTTCAGTGGG	4113
Dd	1		TCAGCTACTCTGGAACCTTGAACTGGAATGAGAGAAATGGGGATRACTTCTCAGTGGG	60
OY	4114		GGAAGGCGAGCTCTTGTGTGATAGCTAGAGACCCCTGCTCCGCCACTGTAAAGTTCTGATTTT	4173
Dd	61		GGAAGGCGAGCTCTTGTGTGATAGCTAGAGACCCCTGCTCCGCCACTGTAAAGTTCTGATTTT	120
OY	4174		AGATGAAGCCACAGCTGCCATGAGACACAGAGACAGACTTAATTGATTTCAAGAACCATCCG	4233
Dd	121		AGATGAAGCCACAGCTGCCATGAGACACAGAGACAGACTTAATTGATTTCAAGAACCATCCG	180
OY	4234		AGAAGCATTTTGAGACTGTACACATGGTGTGACCATTTGGCCATCGCTGCACAGGGTTCTAGG	4293
Dd	181		AGAAGCATTTTGAGACTGTACACATGGTGTGACCATTTGGCCATCGCTGCACAGGGTTCTAGG	240
OY	4294		CTCCGATAGGATTAATGTCGTGCGCCACGAGGACAGGTGTTGAGATTTGACACCCCATCGGT	4353
Dd	241		CTCCGATAGATTTAATGTCGTGCGCCACGAGGACAGGTGTTGAGACCCCATCGGT	300
OY	4354		CCTTTGTGCCAACGACAGCTTCCCGATTCATGCGCATGTTTTGCTGCTGACAGAAACAAGT	4413
Dd	301		CCTTTGTGCCAACGACAGTTCCTCGATTCATGCGCATGTTTTGCTGCTGACAGAAACAAGT	360

FEATURES	SOURCE
BASE COUNT	119 a 150 c 156 g 148 t
ORGANISM	human.
REFERENCE	BE386272 573 bp mRNA EST 21-JUL-2000
AUTHORS	601273486F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3614828 5',
TITLE	mRNA sequence.
JOURNAL	BE386272 BE386272
COMMENT	BE386272.1 GI:9331637
LOCUS	BE386272 573 bp mRNA EST 21-JUL-2000
DEFINITION	601273486F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3614828 5',
ACCESSION	BE386272
VERSION	BE386272.1 GI:9331637
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
TITLE	1 (bases 1 to 573)
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/.
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)
	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.
	Email: cgabs-remail.nih.gov
	Tissue Procurement: ATCC/DC/DPH
	cDNA Library Preparation: Ling Hong/Rubin Laboratory
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
	DNA Sequencing by: Incyte Genomics, Inc.
	Clone distribution: MGC clone distribution information can be
	found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
	Plate: LNCM217 row: 0 column: 21
	High quality sequence stop: 573.
	Location/Qualifiers
	1..573
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:3614828"
	/clone_11d="NIH_MGC_20"
	/tissue_type="melanotic melanoma"
	/lab_host="DH10B (phage-resistant)"
	/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
	EcoRI; cDNA made by oligo-dT priming. Directionally
	cloned into EcoRI/XhoI sites using the following 5'
	adapter: GGCAAGAG(g). Size-selected >500bp for average
	insert size 1.8kb. Library constructed by Ling Hong in
	the laboratory of Gerald M. Rubin (University of
	California, Berkeley) using ZAP-cDNA synthesis kit
	(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 11.8%; Score 570.4; DB 167; Length 573;

Best Local Similarity 99.8%; Pred. No. 1.5e-132; Matches 571; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 2685 TGGCATTCCGTTTATTTATTTAGCCCTTTTCATGCTGTAATGTAGCAGCACCAGCCTTCAGCA 2744
DB 2 TGGCATTCCGTTTATTTATTTAGCCCTTTTCATGCTGTAATGTAGCAGCACCAGCCTTCAGCG 61
OY 2745 CCGTGTGTTGAGTTACTGATCAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2804
DB 62 CCGTGTGTTGAGTTACTGATCAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 121
OY 2805 ACGACACCTCGGTGAGTGCACAGCAGTGAAGGAGCAATCTCATATGCAAGTATGCGCAGCA 2864
DB 122 ACGACACCTCGGTGAGTGCACAGCAGTGAAGGAGCAATCTCATATGCAAGTATGCGCAGCA 181
OY 2865 TGTAGCCCTCCATGAGGAGTGCATGCTGATGAAAGCCATGAGGAGTGTCTTTG 2924
DB 182 TGTAGCCCTCCATGAGGAGTGCATGCTGATGAAAGCCATGAGGAGTGTCTTTG 241
OY 2925 TCAAGGAGCAGCTGCGAGCTTCCTCCGCTGATGACGAGCTTTTCCGAGAGATCCTTC 2984
DB 242 TCAAGGAGCAGCTGCGAGCTTCCTCCGCTGATGACGAGCTTTTCCGAGAGATCCTTC 301
OY 2985 GAGCCCTATGAGTTTATTTTGAACAGACCCCAAGGAGGAGTTCACAGGATTTTCCA 3044
DB 302 GAGCCCTATGAGTTTATTTTGAACAGACCCCAAGGAGGAGTTCACAGGATTTTCCA 361
OY 3045 AAGACATGATTAATGATGACGTCGCGCTGCTTCAGAGCCAGATGTTTCATCCAGAACG 3104
DB 362 AAGACATGATTAATGATGACGTCGCGCTGCTTCAGAGCCAGATGTTTCATCCAGAACG 421
OY 3105 TTAATCGGTGTTCTTCTGCTGAGGAGTATGACGAGAGTCTTCCGCTGCTTCTGAG 3164
DB 422 TTAATCGGTGTTCTTCTGCTGAGGAGTATGACGAGAGTCTTCCGCTGCTTCTGAG 481
OY 3165 CAGTGGGGCCCTTGTATCTCTTTTTCAGTCCGTCACATTTGCTCCAGAGGCTCGATTC 3224
DB 482 CAGTGGGGCCCTTGTATCTCTTTTTCAGTCCGTCACATTTGCTCCAGAGGCTCGATTC 541
OY 3225 GGGAGCTGAAGGCTCTGAGCAATATCAGCGAG 3256
DB 542 GGGAGCTGAAGGCTCTGAGCAATATCAGCGAG 573

```

RESULT 7
 BG403411 679 bp mRNA EST 12-MAR-2001
 DEFINITION 60241908991 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4526110 5',
 mRNA sequence.
 ACCESSION BG403411
 VERSION BG403411.1 GI:13296859
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 679)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: c9abbs-remail.nih.gov
 Tissue Procurement: AFCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.llnl.gov

Plate: LIML10433 row: a column: 23
 High quality sequence stop: 667.
 Location/Qualifiers

FEATURES

1. 679
 source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4526110"
 /clone_lib="NIH_MGC_93"
 /tissue_type="transitional cell papilloma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.7 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC library."

BASE COUNT 157 a 178 c 171 g 173 t
 ORIGIN

Query Match 11.7%; Score 569.2; DB 153; Length 679;

Best Local Similarity 96.8%; Pred. No. 3.2e-132; Matches 666; Conservative 0; Mismatches 13; Indels 9; Gaps 8;

```

OY 3883 GAGATCACTGATATTGGCTTGGCCACTCCGAGCAAACTCTATCTCTCAAGA 3942
DB 1 GAGATCACTGATATTGGCTTGGCCACTCCGAGCAAACTCTATCTCTCAAGA 58
OY 3943 GCCGGTGTCTTCAAGTGGCAGCTGTCAGATCAAAATTTGACCCCTTCACACGATACCTGA 4002
DB 59 GCCGGTGTCTTCAAGTGGCAGCTGTCAGATCAAAATTTGACCCCTTCACACGATACCTGA 117
OY 4003 AACACGATTTGGGATGCCCTGGAGAGACACACATGAAAGATGATGCTCAGCTAC 4062
DB 118 AACACGATTTGGGATGCCCTGGAGAGACACACATGAAAGATGATGCTCAGCTAC 177
OY 4063 TCTGAACCTTGATCTGAGTGAATGAGAGATGGGATTAATCTCTCAGTGGGGAACGGCA 4122
DB 178 TCTGAACCTTGATCTGAGTGAATGAGAGATGGGATTAATCTCTCAGTGGGGAACGGCA 237
OY 4123 GCTCTGTGATAGCTAGAGCCCTGCTCCGCACTGTGAATTTGATTTAGATGAAC 4182
DB 238 GCTCTGTGATAGCTAGAGCCCTGCTCCGCACTGTGAATTTGATTTAGATGAAC 297
OY 4183 CACAGCTGCCATGAGACACAGACAGACTTATTTGATTCAGAGACATCCGAGAAAGATT 4242
DB 298 CACAGCTGCCATGAGACACAGACAGACTTATTTGATTCAGAGACATCCGAGAAAGATT 357
OY 4243 TGCAGACTGTACCATGCTACCATGCCCCTGCGACAGAGTTCTAGGCTCCGATTA 4302
DB 358 TGCAGACTGTACCATGCTACCATGCCCCTGCGACAGAGTTCTAGGCTCCGATTA 416
OY 4303 GATTATGCTGTGGCCAGGAGCAGAGTGTGAGTTGACACCCCAATGCGTCTCTGTC 4362
DB 417 GATTATGCTGTGGCCAGGAGCAGAGTGTGAGTTGACACCCCAATGCGTCTCTGTC 474
OY 4363 CAACGACATTTCCCGATTATGATGCTGCTGCTGTCGACAGAACAGAGTGGCTGCA 4422
DB 475 CAACGACATTTCCCGATTATGATGCTGCTGCTGTCGACAGAACAGAGTGGCTGCA 534
OY 4423 GGGCTGACTCTCCCGATTGAGCAAGTCTTTTCTTTAGAGCATTTGCTCCGCT 4482
DB 535 GGGCTGACTCTCCCGATTGAGCAAGTCTTTTCTTTAGAGCATTTGCTCCGCT 594
OY 4483 GGGGGGGGGCCCTTCAATGCGTCTCTCAAGAAACCTTTCCTTCGATTTATCTTT 4542
DB 595 TGGGGGGGGCCCTTCAATGCGTCTCTCAAGAAACCTTTCCTTCGATTTATCTTT 651
OY 4543 CGCAGACAGCTTCCGATTTGCTGTGT 4570
DB 652 CGCAGACAGCTTCCGATTTGCTGTGT 679

```

RESULT 8

BF816900
 LOCUS BF816900 607 bp mRNA EST 13-JAN-2001
 DEFINITION MR3-C10184-081200-002-f10 C10184 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF816900
 VERSION BF816900.1 GI:12153115
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheraia; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 607)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 TITLE
 JOURNAL
 MEDLINE
 COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL.
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=MR3&l2=MR3-C10184-081200-002-f10&l3=2000-12-08&l4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 588.
 Location/Qualifiers
 1. 607
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_11b="C10184"
 /dev_stage="Adult"
 /note="Organ: colon, ins: Vector: puc18; Site:1: SmaI; Site:2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 131 a 166 c 161 g 148 t 1 others
 ORIGIN

Query Match 11.6%; Score 564.6; DB 169; Length 607;
 Best Local Similarity 98.1%; Pred. No. 4.4e-131;
 Matches 581; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 2797 TCGAGGAGCAGACCTCGTGTGACAGCATGAGCAGCAATCCTCATATGAGTACTA 2856
 |||||||
 Db 4 TCGAGGAGCAGACCTCGTGTGACAGCATGAGCAGCAATCCTCATATGAGTACTA 63
 |||||||
 QY 2857 TGCAGCATACGCCCTCTCCATGAGCAGTATGCTGTAAGGCAATTCAGAGAGT 2916
 |||||||
 Db 64 TGCAGCATACGCCCTCTCCATGAGCAGTATGCTGTAAGGCAATTCAGAGAGT 123
 |||||||
 QY 2917 TGTCTTGTCAAGGCGCTGCGAGCTTCTCCGCGTGCATGACAGGCTTTCCGAG 2976
 |||||||
 Db 124 TGTCTTGTCAAGGCGCTGCGAGCTTCTCCGCGTGCATGACAGGCTTTCCGAG 183
 |||||||
 QY 2977 GATCCTTCGAAGCCCTATGAGTTTGTGACAGACCCCAAGGAGATTCACAG 3036
 |||||||
 Db 184 GATCCTTCGAAGCCCTATGAGTTTGTGACAGACCCCAAGGAGATTCACAG 243
 |||||||
 QY 3037 GTTTCACAAAGACATGATGATGACGTGCGGCTGCCGTTCCAGGCGGAGATGTTTCAT 3096
 |||||||

Db 244 GTTTCACAAAGACATGATGATGACGTGCGGCTGCCGTTCCAGGCGGAGATGTTTCAT 303
 |||||||
 QY 3097 CCAGAACGTTATCCCTGTGTCTTCTGTGTGGAGATGATTCAGAGAGTCTCCGCGGTT 3156
 |||||||
 Db 304 CCAGAACGTTATCCCTGTGTCTTCTGTGTGGAGATGATTCAGAGAGTCTCCGCGGTT 363
 |||||||
 QY 3157 CCTGTGTGCAAGTGGGCCCCCTTGTGCATCTCTTTTCAGTCTGACATCTGTCCAGGCT 3216
 |||||||
 Db 364 CCTGTGTGCAAGTGGGCCCCCTTGTGCATCTCTTTTCAGTCTGACATCTGTCCAGGCT 423
 |||||||
 QY 3217 CCTGTGTGCAAGTGGGCCCCCTTGTGCATCTCTTTTCAGTCTGACATCTGTCCAGGCT 3276
 |||||||
 Db 424 CCTGTGTGCAAGTGGGCCCCCTTGTGCATCTCTTTTCAGTCTGACATCTGTCCAGGCT 483
 |||||||
 QY 3277 CACGTCGACATACAGGCGCTTGGC-ACCATCCAGCGCTTACATTAAGGCGAGAGTTTC 3335
 |||||||
 Db 484 CACGTCGACATACAGGCGCTTGGC-ACCATCCAGCGCTTACATTAAGGCGAGAGTTTC 543
 |||||||
 QY 3336 TGCACAGATACAGGAGCTGCTGATGACACCAAGCTCTTTTGTGT 3387
 |||||||
 Db 544 TGCACAGATACAGGAGCTGCTGATGACACCAAGCTCTTTTGTGT 595
 |||||||

RESULT 9
 BE336596 545 bp mRNA EST 14-JUL-2000
 LOCUS b497c04.y1 NIH.MGC.14 Homo sapiens cDNA clone IMAGE:2957862 5'
 DEFINITION similar to WP:FI4F4.3 CEL1672 ABC TRANSPORTER ; , mRNA sequence.
 ACCESSION BE336596
 VERSION BE336596.1 GI:9188987
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheraia; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 545)
 NIH-MGC http://mhc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Other ESTs: b497c04.x1
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 Image.lnl.gov/image/html/resources.shtml
 Seq primer: -40RP from Gibco
 High quality sequence stop: 453.
 Location/Qualifiers
 1. 545
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2957862"
 /clone_11b="NIH.MGC.14"
 /tissue_type="renal cell adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: kidney; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 126 a 141 c 154 g 123 t 1 others
 ORIGIN

Query Match 11.2%; Score 542.4; DB 166; Length 545;

Best Local Similarity 99.6%; Pred. No. 1.7e-125;
Matches 543; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
OY 162 GGTATGAAAGTGTGAGGGAGAGACACGACATTTCTGGACGACAGAGACCGTGAAGATT 221
    |||||||
DB 1 GGTATGAAAGTGTGAGGGAGAGACACGACATTTCTGGAGAGACAGAGACCGTGAAGATT 60

OY 222 CCAAGTTCGAGGAAGTGTGACGTTGGATGCAAGATGCTTGGAAAGACGACCGCCGAG 281
    |||||||
DB 61 CCAAGTTCGAGGAAGTGTGACGTTGGATGCAAGATGCTTGGAAAGACGACCGCCGAG 120

OY 282 CCGAGGCGCTCTCTCTGATGCGCTCCATGCTATTCAGCTCAGAACTCTGATGAGAGAC 341
    |||||||
DB 121 CCGAGGCGCTCTCTCTGATGCGCTCCATGCTATTCAGCTCAGAACTCTGATGAGAGAC 180

OY 342 ATCCCAAGGAAAGTATACATCATGCTTGTAGTGTCTGAAGCCATCGGACTACTTCCA 401
    |||||||
DB 181 ATCCCAAGGAAAGTATACATCATGCTTGTAGTGTCTGAAGCCATCGGACTACTTCCA 240

OY 402 AACACGACGACCCAGTGTGACATGCTGAGCTTTTTCATGATGCTTTTGTGCTTC 461
    |||||||
DB 241 AACACGACGACCCAGTGTGACATGCTGAGCTTTTTCATGATGCTTTTGTGCTTC 300

OY 462 CTCTCTGCGCCGTGTGCGCCACAAAGAGGGAGCTCTCAATGGAAGACGTGTGCTTC 521
    |||||||
DB 301 CTCTCTGCGCCGTGTGCGCCACAAAGAGGGAGCTCTCAATGGAAGACGTGTGCTTC 360

OY 522 TGTCCAGACAGATCTTTGTGACGTGTAATGCAAGAACTAGAGAGCTGTGCAAGAG 581
    |||||||
DB 361 TGTCCAGACAGATCTTTGTGACGTGTAATGCAAGAACTAGAGAGCTGTGCAAGAG 420

OY 582 ACCTGATGATGAGTGGGCGACAGCGCTTCCTCGAAGGGGTGTGTGATCTTCTGCG 641
    |||||||
DB 421 ACCTGATGATGAGTGGGCGACAGCGCTTCCTCGAAGGGGTGTGTGATCTTCTGCG 480

OY 642 GCACGAGGCTATCTCTGCTCATGCTGTCCTGATGATCAAGCAGCTGCGTTCAGTG 701
    |||||||
DB 481 GCACGAGGCTATCTCTGCTCATGCTGTCCTGATGATCAAGCAGCTGCGTTCAGTG 540

OY 702 GACCA 706
    |||||
DB 541 GACCA 545
```

RESULT 10

```
AJ393794 801 bp mRNA EST 25-JAN-2001
LOCUS AJ393794 dKfz426 Gallus gallus cdna clone 16015r1, mRNA sequence.
DEFINITION AJ393794
ACCESSION AJ393794
VERSION AJ393794.1 GI:7123038
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae;
Phasianidae; Gallus.
REFERENCE 1 (bases 1 to 801)
AUTHORS Abdulkhanov,I., Lodgjin,D., Geroth,P., Arakawa,H., Law,A., Plachy
,J., Korn,B. and Buerstedde,J.M.
TITLE A large database of chicken bursal ESTs as a resource for the
analysis of vertebrate gene function
JOURNAL Genome Res. 10 (12), 2062-2069 (2000)
MEDLINE 20568495
COMMENT Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinstr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES
source
1..801
/organism="Gallus gallus"
/strain="CB"
/db_xref="taxon:9031"
```

/clone="16015r1"
/clone.lib="dKfz426"
/tissue_type="Bursa of Fabricius"
BASE COUNT 225 a 174 c 204 g 198 t
ORIGIN

Query Match 10.6%; Score 511.4; DB 104; Length 801;
Best Local Similarity 80.2%; Pred. No. 1.2e-117;
Matches 626; Conservative 0; Mismatches 151; Indels 4; Gaps 2;

```
OY 3653 AACAGGCTCCCTCCCTCCTACTGCGCCCGAGAGGAGAGAGTGTGAGTGAAGACGAG 3712
    ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 AATTAAGTCCCTCTCTGAGTGTGAGGAGAGAGAGAGTGTGAGTGAAGTGAAGTGAAG 60

OY 3713 ATGAGTACCGAGAAAGTCCCTCTGCTCTAAGAAAGTATCTTCACGATCAAACT 3772
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 ATGCGGTACCGAGAAAGTCCCTCTGCTCTAAGAAAGTATCTTCACGATCAAACT 120

OY 3773 AAAGAAAGTGTGAGTGTGAGGAGGAGAGATCAAGGAAAGTCCCTGAGGAGTGGG 3832
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 AAGAAAGATGTGAGTGTGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180

OY 3833 CTCTTCCTGCTGTGAGTGTGAGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAG 3892
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 CTCTTCCTGCTGTGAGTGTGAGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAG 240

OY 3893 GATATGTGCTTGTGCGAGCTCCGAGAGCAAACTCTATATCTTCTTAAGAGCGGTGCTG 3952
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 241 GATATGTGCTTGTGCGAGCTCCGAGAGCAAACTCTATATCTTCTTAAGAGCGGTGCTG 300

OY 3953 TTCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 4012
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 301 TTCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 360

OY 4013 TGGAGTGCCTGTGAGAGAGAGACACATGAAGAATGTATGTCTGACGTCTGAAACT 4072
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 361 TGGAGTGCCTGTGAGAGAGAGACACATGAAGAATGTATGTCTGACGTCTGAAACT 420

OY 4073 GAATCTGAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4132
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 421 GACTCAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480

OY 4133 ATAGCTAGAGCCCTGCTCCGACCTGATGATGATGATGATGATGATGATGATGATG 4192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 481 ATAGCTAGAGCCCTGCTCCGACCTGATGATGATGATGATGATGATGATGATGATG 540

OY 4193 ATGACACAGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 541 ATGACACAGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 600

OY 4253 ACCATGCTGACATGCTCCGACCTGACACAGCTTCTAGAGCTCCGATAGATGATG 4312
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 601 ACTATGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660

OY 4313 CTGGCCAGAGAGAGTGTGAGAGTGTGACACCCCATCGCTCTTCT-GTCCAAACAGAG 4371
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 661 CTAAACACAGAGAGAGTGTGAGAGTGTGACACCCCATCGCTCTTCTGAGATGAG 720

OY 4372 TTCCCATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4428
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 721 CTGGCCCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780

OY 4429 A 4429
    |
DB 781 A 781
```

RESULT 11
AI023781 495 bp mRNA EST 28-AUG-1998
LOCUS AI023781/c
DEFINITION ox08b10.x1 Soares fetal liver spleen, INFUS.S1 Homo sapiens cDNA
clone IMAGE:1655707 3' similar to SW.Y115_YEAST P14772 PROBABLE

Accession	APF-DEPENDENT TRANSPORTER YLL015w. [1] ; mRNA sequence.
VERSION	AJ023781
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (infoimage.lnl.gov) for further information. Insert Length: 663 Scd Error: 0.00 Seq primer: -40ml3 fwd. ET from Amersham High quality sequence stop: 442.
TITLE	Location/Qualifiers
FEATURES	1..495
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606" /clone IMAGE:1635707" /clone_id="Soares_fetal_liver_spleen_INFLS-S1" /sex="male" /dev_stage="20 week-post conception fetus" /lab_host="DH10B (ampicillin resistant)" /note="Organ: Liver and Spleen; Vector: pTZ19AD (Pharmacina) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI. This is a subtracted version of the original Soares fetal liver spleen INFLS library. 1st strand cDNA was primed with a Pac I - o1igo(dT) primer [5'] AACGGAAGAATTTAATAAGATCTTTTTTTTTTTTTTTT 3'] (Pharmacina), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pTZ19 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT	130 a 120 c 148 g 97 t
ORIGIN	
Query Match	9.9%; Score 478.4; DB 15; Length 495;
Best Local Similarity	99.8%; Pred. No. 2.2e-109;
Matches 479; Conservative	0; Mismatches 1; Indels 0; Gaps 0.
OY	2836 CAATCCTCATATGCAGTACTATGCCAGCATTCACGCCCTCCTCATGGCAGTCATGTGAT 2895
DB	435 CAATCCTCATATGCAGTACTATGCCAGCATTCACGCCCTCCTCATGGCAGTCATGTGAT 436
OY	2896 CCTGAAGCCATTGCAGAGATTGCTTTGTCAAGGGCAGCTGCAGCTTCTCCGGCT 2955
DB	435 CCTGAAGCCATTGCAGAGATTGCTTTGTCAAGGGCAGCTGCAGCTTCTCCGGCT 376
OY	2956 GCATACAGCAGCTTTCCGGAAGATCTTGGAAGCCCTATGAAAGTTTGTGACAACGACCCC 3015
DB	375 GCATACAGCAGCTTTCCGGAAGATCTTGGAAGCCCTATGAAAGTTTGTGACAACGACCCC 316
OY	3016 CACAGGAGAGATTTCACAGAGCTTTTCCAAACACATGATGTAAGTTGACGTCGGCTGCC 3075
DB	315 CACAGGAGAGATTTCACAGAGCTTTTCCAAACACATGATGTAAGTTGACGTCGGCTGCC 256
OY	3076 GTTCCAGGCGCAGATGTTTCATCCAGAACGTATCTCGGTCTTCTGTGTGGGAATGAT 3135
DB	255 GTTCCAGGCGCAGATGTTTCATCCAGAACGTATCTCGGTCTTCTGTGTGGGAATGAT 196
OY	3136 CGCAGAGAGCTTCCGCTGCTTCTGTGGCAGTGGGGGCCCTGTGATCCTCTTTTCAGT 3195
DB	195 CGCAGAGAGCTTCCGCTGCTTCTGTGGCAGTGGGGGCCCTGTGATCCTCTTTTCAGT 136
OY	3196 CCGTGCACATTTGCTCCAGGCTCTGATTTCCGGAGCTGAAGCGCTGTGACAAATATCACGA 3255

Db	135	CCTGCACATTTCCTCCAGGCGTCTGATTCGGAGACTGAAGCGTCTGACAAATATCACCA	76
QY	3256	GTCACCTTTCTCTCCTCCACATCATAGCTCCAGCATATCAGGGCCTTGCCACCATCCAGCCTTA	3315
Db	75	GTCACCTTTCTCTCCTCCACATCATAGCTCCAGCATATCAGGGCCTTGCCACCATCCAGCCTTA	16
RESULT	12		
LOCUS	BE754240	520 bp	mRNA EST 15-SEP-2000
DEFINITION	207698 MARC 2Bov Bos taurus CDNA 5', mRNA sequence.		
ACCESSION	BE754240		
VERSION	BE754240.1	GI:10168232	
KEYWORDS	EST.		
SOURCE	cow.		
ORGANISM	Bos taurus		
	Eumaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
	Bovidae; Bovinae; Bos.		
	1 (bases 1 to 520)		
REFERENCE	Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,		
AUTHORS	Bennett,G.A., Fahrnkung,S.C., Freking,B.A., Rohrer,G.A., Laegreid		
	,W.W. and Keele,J.W.		
TITLE	Design and use of four pooled tissue normalized cDNA libraries for		
	EST discovery in cattle		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Smith TPL		
	USDA, ARS, US Meat Animal Research Center		
	PO Box 166, Clay Center, NE 68933-0166, USA		
	TEL: 402 762 4366		
	Fax: 402 762 4390		
	Email: smithhemall.marc.usda.gov		
	Single pass sequencing. Bases called and alt-trimmed with phred		
	v0.980904.e. Vector identified by cross_match with the -minscore 18		
	and -mismatch 12 options.		
	PCR primers		
	FORWARD: AGGAACAGCATATGACCAT		
	BACKWARD: GTTTGCCAGTCCAGCAGC		
	Plate: 51 row: A column: 22		
FEATURES	Seq primer: ATTTAGGTGACATATAG.		
source	Location/Qualifiers		
	1..520		
	/organism="Bos taurus"		
	/db_xref="taxon:9913"		
	/clone_lib="MARC 2BOV"		
	/tissue_type="pooled"		
	/lab_host="DH10B"		
	/note="Vector: pCMV SPORT6; Site_1: XbaI; site_2: XhoI;		
	library made from pooled tissue from testis, thymus,		
	semilendinous muscle, longissimus muscle, pancreas,		
	adrenal, and endometrium."		
BASE COUNT	126 a 140 c 132 g 122 t		
ORIGIN			
Query Match	9.6%; Score 465.6; DB 139; Length 520;		
Best Local Similarity	93.5%; Pred. No. 3,7e+06;		
Matches	486; Conservative 0; Mismatches 34; Indels 0; Gaps 0;		
QY	3882	TGAGATCATGATGATTTGGCCTTGGCGACCTCCGAAGCAAACTCTGTATCATTTCTCTCAAG	3941
Db	1	TGAGATCATGATGATTTGGCCTTGGCGACCTCCGAAGCAAACTCTGTATCATTTCTCTCAAG	60
QY	3942	AGCCGCTGCTGTTCACTGGCACTGTGCATGATTCAAATTTGGACCCCTTCAACAGGTACAGT	4001
Db	61	AGCCGCTGCTGTTCACTGGCACTGTGCATGATTCAAATTTGGATCCCTTCAACAGGTACAGT	120
QY	4002	AGACACAGATTTGGAGTGGCCTTGAGAGAGACACACATGAAAAGATGTATTGCTCAGCTAC	4061
Db	121	AGACACAGATTTGGAGTGGCCTTGAGAGAGACACACATGAAAAGATGTATTGCTCAGCTAC	180
QY	4062	CTCTGAACCTGATTCGATGAGATGAGACAATGGGATTAATCTTCTCAGTGGGGAGAACGGC	4121


```

Db 418 TCCTGGTATCATGTCCTCTTCATGCTGAATGTGGGACAGACACTGCTTCAGACACTGCT 477
OY 2751 GGTGAGTTACTGATCAAGCAAGGAGGAGACACCACTGTGACTGAGGGAACGAGA 2810
Db 478 GCTTATCTACTGACCAAGCAAGGAGGAGACAGCAGCAGTGTATCAATGGAACAGAA 537
OY 2811 CCTGGTGTGACAGCANTGAAGCAATCCCTCATATGCACTACTA 2856
Db 538 NCTTCGTGAGTGACAGCANTGAATGACAGCCCTTCATGCACTACTA 583

```

Search completed: November 9, 2001, 14:03:15
 Job time: 8563 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: November 9, 2001, 15:51:42 ; Search time 73.13 Seconds
(without alignments)
1191.258 Million cell updates/sec

Title: US-09-528-031-2

Perfect score: 7308

Sequence: 1 MKDDIGREYIIIPSPGYRSV.....DSSRFYAMFAAENKVAWG 1437

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_0601:*

1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:*

2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:*

3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:*

4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:*

5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:*

6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:*

7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:*

8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:*

9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:*

10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:*

11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:*

12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:*

13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:*

14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:*

15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:*

16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:*

17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:*

18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:*

19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:*

20: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:*

21: /SIDSL/gcgdata/geneseq/geneseq/AA2001.DAT:*

22: /SIDSL/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7308	100.0	1437	21	AAAB10225
2	7304	99.9	1437	19	AAAB0357
3	7298	99.9	1437	22	AAAB7021
4	7293	99.8	1437	20	AAV43542
5	2094.5	28.7	1541	18	AAW33361
6	2082	28.5	1531	19	AAW57486
7	2082	28.5	1531	20	AAW99894
8	2082	28.5	1531	20	AAW74471
9	2082	28.5	1531	21	AAV78873
10	2082	28.5	1531	21	AAV55799
11	2082	28.5	1545	18	AAW33362

12	2075.5	28.4	1489	17	AAW6952	Multi-drug resista
13	2075	28.4	1531	15	AAW54928	Multi-drug resistan
14	2075	28.4	1531	17	AAW3153	Multi-drug resista
15	2075	28.4	1531	17	AAW57485	Human multidrug re
16	2075	28.4	1531	20	AAW99893	Human multidrug re
17	2075	28.4	1531	20	AAW74470	Human multidrug re
18	2075	28.4	1531	21	AAW3582	Multi-drug resistan
19	2075	28.4	1531	21	AAV78872	Human multidrug re
20	2075	28.4	1531	21	AAV5798	Human multidrug re
21	2070	28.3	1545	19	AAW55966	Human canalicular
22	2055.5	28.1	1528	19	AAW57487	Murine multidrug r
23	2055.5	28.1	1528	20	AAW99895	Mouse multidrug re
24	2055.5	28.1	1528	20	AAW74472	Mouse multidrug re
25	2055.5	28.1	1528	21	AAV78874	Murine multidrug r
26	2055.5	28.1	1528	21	AAV55800	Murine multidrug r
27	2049.5	28.0	1525	20	AAV3541	A human MPR-relate
28	2006.5	27.5	1482	17	AAW6953	Multi-drug resista
29	1993.5	27.3	1622	19	AAW60163	Arabidopsis thalia
30	1977	27.1	1621	19	AAW60162	Arabidopsis thalia
31	1941.5	26.6	1528	18	AAW33363	Human multidrug re
32	1941	26.6	1527	20	AAV3543	A human MPR-relate
33	1861.5	25.5	1472	17	AAW6954	Multi-drug resista
34	1828	25.0	1355	21	AAW3084	Arabidopsis thalia
35	1807.5	24.7	1352	21	AAW6261	Arabidopsis thalia
36	1807.5	24.7	1511	21	AAW6260	Arabidopsis thalia
37	1778.5	24.3	1545	19	AAW53602	Rat sulphuryurea
38	1769.5	24.2	1215	21	AAW30842	Arabidopsis thalia
39	1755	24.0	1147	21	AAW30843	Arabidopsis thalia
40	1749	23.9	1212	21	AAW6262	Arabidopsis thalia
41	1745.5	23.9	1580	19	AAW57412	Homo sapiens sulph
42	1743	23.9	1772	21	AAW52485	Arabidopsis thalia
43	1743	23.9	1776	21	AAW52484	Arabidopsis thalia
44	1743	23.9	1787	21	AAW52483	Arabidopsis thalia
45	1730	23.7	1477	18	AAW10424	Saccharomyces cere

ALIGNMENTS

RESULT	1
AAAB10225	standard; Protein: 1437 AA.
ID	AAAB10225
AC	AAAB10225;
DT	16-NOV-2000 (first entry)
DE	Human MRP-beta protein.
XX	
KW	MRP-beta; multidrug-resistance associated protein; human; cytostatic;
KW	chemotherapy; cancer cell; drug discovery; cytotoxic.
XX	
OS	Homo sapiens.
XX	
PN	US6077936-A.
XX	
PD	20-JUN-2000.
XX	
PF	16-APR-1998; 98US-0061400.
XX	
PR	16-APR-1997; 97US-0843459.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
PI	Shyjan A;
XX	
DR	WPI: 2000-430613/37.
XX	
DR	N-PSDB: AAA40481.
XX	
PT	Model multidrug-resistance associated polypeptide useful for improving
PT	the effectiveness of a chemotherapeutic regimen to eradicate
PT	multidrug-resistant transformed cells especially cancer cells -

PS Claim 1, Fig 1A-D; 43pp; English.
XX
CC This invention describes a novel model multidrug-resistance associated
CC polypeptide, MRP-beta (1) which has cytostatic activity. Inhibitors of
CC (1) are useful for improving the effectiveness of a chemotherapeutic
CC regimen to eradicate multidrug-resistant transformed cells, especially
CC cancer cells, from the body of a mammal, preferably human. (1) is also
CC useful for drug discovery, especially to the design of novel
CC chemotherapeutic drugs that are cytotoxic to cells expressing (1). This
CC sequence represents the human MRP-beta protein which is described in the
CC method of the invention.
XX
SQ Sequence 1437 AA:

Query Match 100.0%; Score 7308; DB 21; Length 1437;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDIDIGKEITIPSGYRSVRETSSTGTHREDSEKFRRTREPCODALETAARAGLS 60
DB 1 MKDIDIGKEITIPSGYRSVRETSSTGTHREDSEKFRRTREPCODALETAARAGLS 60
QY 61 LQASMSQRLIDEBEHPKGYHNGLSAKPIRTTCQHGVNNAAGLFSCMTFWSLSLAR 120
DB 61 LQASMSQRLIDEBEHPKGYHNGLSAKPIRTTCQHGVNNAAGLFSCMTFWSLSLAR 120
QY 121 VARKKGLMEDWMSLSKSHSSDVNCRRLERLMOEELNENYGPDAASLRVWVICRRLI 180
DB 121 VARKKGLMEDWMSLSKSHSSDVNCRRLERLMOEELNENYGPDAASLRVWVICRRLI 180
QY 181 LSTVICMTTOLAGFSGPAFWVKHLEVTQATESNLQYSLLVLCGLLTETIVRSWLSLTW 240
DB 181 LSTVICMTTOLAGFSGPAFWVKHLEVTQATESNLQYSLLVLCGLLTETIVRSWLSLTW 240
QY 241 ALMYRGSVRLRGALITMAFKKILKLNKEKSGELINICNSDQRMFEAAVSSLAGG 300
DB 241 ALMYRGSVRLRGALITMAFKKILKLNKEKSGELINICNSDQRMFEAAVSSLAGG 300
QY 301 PVAALIGMIVNIIGPTGFLGSAVFLEFYPAWMAFASRLRAYFRKCVAAATDERVOKME 360
DB 301 PVAALIGMIVNIIGPTGFLGSAVFLEFYPAWMAFASRLRAYFRKCVAAATDERVOKME 360
QY 361 VLYYIKRKMYAWKAFSOSVOKIREBERILEKAGYFOSITGVAPIVVIAVSVFVS 420
DB 361 VLYYIKRKMYAWKAFSOSVOKIREBERILEKAGYFOSITGVAPIVVIAVSVFVS 420
QY 421 HMTLGEPUAAQAFVTVTVNSMTFALKVTFPSYKSLSEASVAVDRKSLFLMEVMIK 480
DB 421 HMTLGEPUAAQAFVTVTVNSMTFALKVTFPSYKSLSEASVAVDRKSLFLMEVMIK 480
QY 481 NKRPASPIKTEKKNATLAWPSSSHSSIONSPLPKMKKDRASGKKEKRYOLORTHEOA 540
DB 481 NKRPASPIKTEKKNATLAWPSSSHSSIONSPLPKMKKDRASGKKEKRYOLORTHEOA 540
QY 541 VLAEGQKHLILDSDERPSEEEGKNIHLGLRLQRTLSIDLEIQGKLVGICGSVSG 600
DB 541 VLAEGQKHLILDSDERPSEEEGKNIHLGLRLQRTLSIDLEIQGKLVGICGSVSG 600
QY 601 KTSLSIALIGQMTLSESTASISGFAYVAQAAWTLNLTLDNLLFGKEYDEERNSVLNS 660
DB 601 KTSLSIALIGQMTLSESTASISGFAYVAQAAWTLNLTLDNLLFGKEYDEERNSVLNS 660
QY 661 CCLRPDLAIPSSDLTEIGEGANLSCGORISLARALYSDRSITLIDPLSLDAHV 720
DB 661 CCLRPDLAIPSSDLTEIGEGANLSCGORISLARALYSDRSITLIDPLSLDAHV 720
QY 721 NHTFNSAIRKHLKSTVLFWTLOLVDCDEVLFMKEGCITERGTHEELMNLNGDYATI 780
DB 721 NHTFNSAIRKHLKSTVLFWTLOLVDCDEVLFMKEGCITERGTHEELMNLNGDYATI 780
QY 781 FNNLLGETFPVELNSKKEKSGSKSQDQKPKTGSITKEKAVKPEEGQLVLEKQGS 840
DB 781 FNNLLGETFPVELNSKKEKSGSKSQDQKPKTGSITKEKAVKPEEGQLVLEKQGS 840

DB 781 fnnllgetfpvelnskkesgsksqdkgpktsiktkekavkpeegqlvleekqgs 840
QY 841 VPMVSYGVTYIOAGCPALFVITMALFMLANVSTAFSTWMLSTYKIOGSGNTTYRGNETS 900
DB 841 VPMVSYGVTYIOAGCPALFVITMALFMLANVSTAFSTWMLSTYKIOGSGNTTYRGNETS 900
QY 901 VSDSMKQNPMMQYVYASITAYALSMAMVLLKAIKGVYFVKGLIRASSKRLHDELFRILLSPM 960
DB 901 VSDSMKQNPMMQYVYASITAYALSMAMVLLKAIKGVYFVKGLIRASSKRLHDELFRILLSPM 960
QY 961 KEFDTPTGRLINRFSKDMQEVDRLPQOAEFMFIONTILVFPFCGMITAGVPMFLVAVGP 1020
DB 961 KEFDTPTGRLINRFSKDMQEVDRLPQOAEFMFIONTILVFPFCGMITAGVPMFLVAVGP 1020
QY 1021 LVILFVLIHIVSVLIRELRLKLDINTQSPFLSHITSSIOGLATTIHAANKGOEFLHRYOEL 1080
DB 1021 LVILFVLIHIVSVLIRELRLKLDINTQSPFLSHITSSIOGLATTIHAANKGOEFLHRYOEL 1080
QY 1081 LDDNQAPPELFTCAMRWLAVRDLISALTITTTGLMTVLMHGQIPPAYAGLAISYAVQLT 1140
DB 1081 LDDNQAPPELFTCAMRWLAVRDLISALTITTTGLMTVLMHGQIPPAYAGLAISYAVQLT 1140
QY 1141 GLPFTYRLASEMFAEFTSVIRINHYIKTSLSEAPARITKNKAPSPDPQGEVTFEAAEM 1200
DB 1141 GLPFTYRLASEMFAEFTSVIRINHYIKTSLSEAPARITKNKAPSPDPQGEVTFEAAEM 1200
QY 1201 RYRENPLVLVTKVYFTIKPEKIGIVRTGSGKSSLGMAFLRVELSGGCIKIDGVRI 1260
DB 1201 RYRENPLVLVTKVYFTIKPEKIGIVRTGSGKSSLGMAFLRVELSGGCIKIDGVRI 1260
QY 1261 IGLADLRSLIIPQEVYLFSGTVRSNLDPPNOYTEDQIWDALERTIMKECIAOLPLKLE 1320
DB 1261 IGLADLRSLIIPQEVYLFSGTVRSNLDPPNOYTEDQIWDALERTIMKECIAOLPLKLE 1320
QY 1321 SEVENGDNSVSGEROLICTARALLRHCKILLIDEATAADTEFDLLIOETIRAFADCT 1380
DB 1321 SEVENGDNSVSGEROLICTARALLRHCKILLIDEATAADTEFDLLIOETIRAFADCT 1380
QY 1381 MLTFAHRLHTVIGSDRLMVLACQOVVEFDPSVLLSNDSSRFYAMFAAENKVAVK 1437
DB 1381 MLTFAHRLHTVIGSDRLMVLACQOVVEFDPSVLLSNDSSRFYAMFAAENKVAVK 1437

RESULT 2
ID AAM80597 standard; Protein: 1437 AA.
XX
AC AAM80597;
XX
DT 02-FEB-1999 (first entry)
XX
DE Human multidrug resistance-associated protein.
XX
KW Multidrug resistance-associated protein; MRP-beta; human; cancer;
KW tumour; chemotherapy; therapy.
XX
OS Homo sapiens.
OS
PN WO9846736-A1.
PN
PD 22-OCT-1998.
PD
PF 16-APR-1998; 98WO-0507673.
PF
PR 16-APR-1997; 97US-0843459.
PR
PA (MILL-) MILLENNIUM PHARM INC.
PA
XX Shyjan A;
XX
PI
PI
DR WPI: 1998-568724/48.
DR N-PSDB; AAV65682.
XX

PT New isolated multidrug resistance-associated polypeptide - used to
PT develop products for modulating multidrug resistance, particularly
PT for reducing resistance of tumours to chemotherapeutic drugs

PS Claim 26; Page 54-59; 93pp; English.

XX This is the amino acid sequence of novel human multidrug resistance
CC associated protein (MRP-beta), as deduced from an isolated cDNA
CC clone (see AAY65682). Over-expression of MRP-beta is thought to be
CC associated with the emergence and/or persistence of a multidrug-
CC resistance phenotype in transformed mammalian cells, including
CC carcinoma and adenocarcinoma cells. MRP-beta is probably a
CC transmembrane-spanning, energy-dependent transporter or pump that
CC functions by ejecting intracellular substances. The invention
CC provides compositions and methods for improving the effectiveness
CC of chemotherapeutic regimens to eradicate multidrug-resistant
CC transformed cells from the body of a mammal, especially from the
CC body of a human. The disclosed compositions include MRP-beta
CC nucleic acids, including probes and antisense oligonucleotides
CC (see also AAY65684-88). MRP-beta polypeptides and antibodies,
CC MRP-beta expressing host cells, and non-human mammals that are
CC transgenic or nullizygous for MRP-beta. The disclosed methods
CC include methods for attenuating aberrant MRP-beta gene expression,
CC protein production and/or protein function, and for improving the
CC effectiveness of chemotherapy for a mammal afflicted with a
CC multidrug-resistant tumour, wherein the tumour is of mammary,
CC respiratory tract, urogenital tract, endocrine system or immune
CC system origin. In addition, methods are disclosed for identifying
CC and using a modulator, such as an inhibitor, that is cytotoxic to
CC cells expressing MRP-beta.

XX Sequence 1437 AA:

Query Match 99.9%; Score 7304; DB 19; Length 1437;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDIDIGEYIIPSPGYSVRERTSTSGTHRDREDSKFRFRPLECODALETAAEGIS 60
DB 1 mkddidigeyiipspgysvrvrtstsgthrdredskfrfrplecodaletaaeagls 60
QY 61 LDASMSQRLRIDEBHPKGYHNGSLAKPIRTTCKHOHPVONAGLFCOMPTSMISLAR 120
DB 61 ldsmsqrlriidebhpkyhngslakpirttckhophvona glfscmtswisslar 120
QY 121 VAHKKGELMEDVWSLSKRESSDVNCRRLERLMQELNEVGDDASLRVYVIFCRTL 180
DB 121 vahkkgelmedvwslskressdvncrrlerlmqelnevgdda slrvvwi fctrl 180
QY 181 LSIYGLMTTQLAGSPGAFNPKHLELYQATFSNIQYSLILVGLLLEIYVMSLALTW 240
DB 181 lsiyglmttqlagspgafnmkhlleyqatesniqysllvlgllleivvmslaltw 240
QY 241 ALNRYTGVRGLGAILTMAFKKILKLNKEKSLGELINCSNDGQRMFEAAVGSILAG 300
DB 241 alnrytgrlrgailtmafkilklnkelslgelincsndgqrmfeaaavgsilagg 300
QY 301 PVVALIGMIVYVILIGPGYLGSAVILFYPAAMPASRLTAYFRRKCYAATDERVOKME 360
DB 301 pvvaligmivvyligpgylgsavilfypammpasrltayfrkrckyaatdervokme 360
QY 361 VLTYYIKFKMTAWKAFESQSQKTEEBRRILIEKAGYFQSTITVGAPYVYVYASVYTSV 420
DB 361 vlttyikfkmtawkafesqsokteebrriliekagyfqsitlvgapivvyvlasvvtlsv 420
QY 421 HMTLGFDTLTAQAFTVYVFNFSMTFALKVTFPSYKSLSEASAVADVRFKSLFMEEVHMK 480
DB 421 hmtlgfdtltaaqftvyvfnfsmfalkvtfpsvkslseasavadvrfkslfmeevhmk 480
QY 481 NKPASPHRIEKKNATLAWDSHSSIQNSPKITPRMKDKRASKCKEKERVOLQRTFHOA 540
DB 481 nkpasphrikknatlawdshssiqnspkltprmkdkraskckekervqlqrtfhoq 540

QY 541 VLAEOGHLIDSDRSPSEEEGKHILGHLRLQRTLHSIDLEIOEGLVIGSGVSG 600
DB 541 vlaeoghli d sdrspseee gkhilghlrlqrtlhsidleioeglvigsgvsg 600
QY 601 KTSLSIALIGQWTLLEGSLAISGTFAYVAOQAMINATLRDNLGKKEYDERYSVLNS 660
DB 601 ktslsialigqwtllegslaisgtfayvaqawlnatlr dnl gkkeyde r ysvlns 660
QY 661 CCLRPDLAILESSDLTEIGERCANISGGORORISARALYSRSTIYLDDPLSALDAHVG 720
DB 661 cclrpdlail essdlteiger canisggororisaralysrstiyl ddp lsaldahvg 720
QY 721 NHIFSAIRKHLKSTVLFVTHQLOLYDCDEVITMKKECITERGTHETMLNNGDYATI 780
DB 721 nhifsa irkhlkstvlfvthqlo lydcdevitmkkec iter gthetmlnngdyati 780
QY 781 FNNLLIGETPVEINSKETSQSKRSODKPKTGSIRKEKAVKEEGQLVLEEKGGQS 840
DB 781 fnnlligetpveinsket s qskrsodkpk tgsirke kavkeegqlvleekgqs 840
QY 841 VPMVYGYVYIQAGGPLAFVLVIMALFMLNVGSTABSTWMLSTWIKQSGNTTVTGNETS 900
DB 841 vpmvygyvyiqaggp l afvlvimalfmlnvgstabstwm l stw i kqsgnt tv tgnets 900
QY 901 VSDSKDNPHMQYASISYALSMANVMLIKAIRGVYFVGTLTAASSRLDELFRRLRSPM 960
DB 901 vsdskdnphmqyas is yal smavmlikairgv yfvgtlta assrl del fr r l r spm 960
QY 961 KPFDTTPGRIINRFSKMDDEVVRLPFOAEMFIQNVILVFCVCMINGVFPFPLVAVGP 1020
DB 961 kpfdttpgr i inrfskmddevvrlpfoaemfiqnvilvfcvcm ingvfpfplvavgp 1020
QY 1021 IVILFVSILVSRVLLIRELKRLDNTQSPFLSHITSIOGLATIAHYNGQEFRLRYQEL 1080
DB 1021 ivilfvsilv sr vllirelkrldntqspflshitsio glat i ahyngqefrlryqel 1080
QY 1081 LDDNAPPEFLFCAMRWLAVRLDLSIALITTTGIMLYVMHQIQIPAYAGLISYAVQLT 1140
DB 1081 lddnappeflfcamrw l avrldlsialitttgimlyvmhqiq ipayagla isyavqlt 1140
QY 1141 GLFOFTVRLASETEARFVSVERINHYKITLSLEAPARIKNKAPSPDMPQEGEVTEFNAEM 1200
DB 1141 glf o f t v r l a s e t e a r f v s v e r i n h y k i t l s l e a p a r i k n k a p s p d m p q e g e v t e f n a e m 1200
QY 1201 RYRENLPVLKRVSTFKPKKEKIGIVRTSGKSSLGMAFLPVLDELGGCIIIDGVRISD 1260
DB 1201 ryrenlpvlk r vstfkpkkekigivrtsgksslgma flpvl del ggc i i d g v r i s d 1260
QY 1261 IGLADLRKSLSTIIPQEPVLFSGTAVSNLDPFNQYTEDQIWMALERTHMKECIAQLPKLE 1320
DB 1261 igladlrkslsti ipqepvlfsgtavn l dpfnqytedqiwmal erthmk eciaqlpkle 1320
QY 1321 SEVMENGNDFSVGERQLCIARALRHCKITLIDETATAMDETETLLOETIRREAFADCT 1380
DB 1321 sevmengndfsvgerqlciaralrhckitl idetata m detetlloetirreafadct 1380
QY 1381 MLIARHLTVLGSORINVLAGOVVEEDFVSVLNSDSRFYAMPAAENKVAVKG 1437
DB 1381 mliarhl tvlgsorin v l a g o v v e e d f v s v l n s d s r f y a m p a a e n k v a v k g 1437
RESULT 3
AAB47021
ID AAB47021 standard; Protein: 1453 AA.
XX
AC AAB47021;
XX
DT 29-MAR-2001 (first entry)
XX
DE Multidrug-resistance associated polypeptide-beta.
XX
KW Multidrug resistance-associated polypeptide; MRP-beta; transporter;

KM P-glycoprotein; chemotherapy; cancer.
XX OS Homo sapiens.
XX US6162616-A.
PN XX
XX 19-DEC-2000.
PD XX
XX 16-APR-1997; 970S-0843459.
XX PF
XX 16-APR-1997; 970S-0843459.
PR XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX ShyJan A;
PI
XX WPI: 2001-111728/12.
DR N-PSDB; AAC85287.
PS
PT Novel nucleic acid encoding multidrug resistance-associated protein
PT MRP-beta, useful for detecting and treating drug-resistant cancers -
XX
PS Disclosure; Fig 1; 41pp; English.
XX
CC This sequence represents a multidrug resistance-associated
CC polypeptide (MRP-beta). MRP-beta is a transporter that can remove
CC chemotherapeutic agents from cells and/or sequester such agents,
CC and it probably accounts for multidrug resistant phenotypes that do
CC not respond to treatments with reversal agents directed against
CC P-glycoprotein or the known MRP. By inactivating MRP-beta, at nucleic
CC acid or protein levels, cells should be rendered sensitive to
CC chemotherapy agents. The MRP-beta coding sequence, or fragments
CC of it, can be used to detect expression of the MRP-beta gene and
CC thus multidrug resistant cells. Fragments of MRP-beta DNA are useful
CC as primers and as antisense (therapeutic) agents for inhibiting
CC expression of the MRP-beta gene. They can also be used for detecting
CC mutations and to treat diseases associated with abnormal MRP-beta
CC gene expression, specifically cancer, particularly for potentiating
CC chemotherapy of fluid or inoperable tumors, or in cases of regrowth
CC after an initial course of chemotherapy.
CC
XX
XX Sequence 1453 AA:
SQ
Query Match 99.9%; Score 7298; DB 22; Length 1453;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1436; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKDDIDKEVYIIPSPGYSRVERKSTSGTHRDSDSKFRKRRPLECDALETARAGLS 60
DB 17 mkddidkeylilpspyrsvrertstsgthrdedskfrrtrplecdaletaaraegls 76
QY 61 LDASMHQSLRLDDEHKGKYNHGLSALPKRTCKHOHPDNAGLFCSCMFTSLSLAR 120
DB 77 ldsamhbsqllildehnpkyknhnglsalpkrrttsknhpndnagllfscmftslslar 136
QY 121 VAHKGELSMEDVWSLSKHSSSDVNCRLERLMOBELNEVGPDAASLRVWICRRLI 180
DB 137 vahkgelmedvwslskhsdsvncrrlerlmoelnevgsdaaslrwvicrrtll 196
QY 181 LSTVCLMITQLAGSGAFAMWKHLLETQATSNLQYSLLVGLLLETYSRSLATW 240
DB 197 lsvclmitqlagsgafamwkhlletqatenslqysllvlgllletyrsvaltw 256
QY 241 ALNRTGVRRLGATLTAAFKILTKNIKESLGLINICSDGCRFEAAVAGSLAAG 300
DB 257 alnrtgvrirlgatltaafkilkiknikeslgellincsdgrmfcaavagslaag 316
QY 301 PVVAILGMIVNIIIGPTGLGSAVFTLFFPAMMFASRLTAYFRKCAATDERVOKNE 360
DB 317 pvvailgmivniiigptglgsavftlffpammfasrltayfrkcaatdevvokne 376
QY 361 VLYIKIKIMYAWKAFSOSVOKITREBRILKEAGYFQSTTGVAPIVVIVASVTF 420

DB 377 VLYIKIKIMYAWKATSGVQKIREERRILEKAGYFGSTLVGVAPIVVIVASVTF 436
QY 421 HMTLGFDTAAQAFVTVTNVNSMTFALKVTPFSYKSLSEASVAVDRKSLFLEMEVMIK 480
DB 437 hmtlghdtaaqafvtvtvntvnsmtfalkvtpfsykslseasvavdrkslfleemv 496
QY 481 NKPASPIKTIEMKNATLAMDSSSIONSPKLTVPKMKDKRASGKKKRYOLORTTHOA 540
DB 497 nkpaspiikiemknatlamdssshionspkltvpkdkdkrasgkkrkryolrthoa 556
QY 541 VLAOKGHLLDSDRSPPEEKGKHLHGLRLQRTLSHIDLEIOGKLVGTGSGVSG 600
DB 557 vlaeqlhllsdersppeeegkhlhglrlqrlshidlelgeqklvgjcsqvsg 616
QY 601 KTSLSIALGQMTLEGSIAISGTFAYVAAOAWILNATRLDNILFGEYDEERYNSYLS 660
DB 617 ktslslalgqmtllegsiaisgtfayvaagawilnatrlndnllfgyeeydeeryns 676
QY 661 CCLRPDAILPSSDLTEIGERGANLSCGQORISLARALYSDRSITYLLDPLSLADHVG 720
DB 677 cclrpdaillpsdlteligeranlsgqqrtslaralysdrsiyllddplsaldahv 736
QY 721 NHTNSAIRKHLKSKTVLPTVHOLOYVDCDEVTFMKEGCTTERGTHEELMNLNGVATI 780
DB 737 nhlnsairkhlksktrlvptvhoqlyvdcdevlfmkegcttergtheelmnlngvati 796
QY 781 FNNILLGETPPEVENSKRKETSQSKSODKPKTSIRKEKAVRBEQOLVOLBEKQGS 840
DB 797 fnnillgetpvevnskrsqsksgkdkpkslkkkavpbeeqqlvgleekqgs 856
QY 841 VPMVSYGYVIOAAGPLAFVYLMALFNLNNGSTAFSTWMLSYWKQSGNTVYRGETS 900
DB 857 vpmvsygyvioaagplafvylmalfnlmgstafstwmlsywkqsgntvlyrgets 916
QY 901 VSDMKNPBHQYVASYALSMAYMLIKATRGVFPKGTLRASSRLHDELFRLILSPM 960
DB 917 vsdmknbpbmqyasyalasmaaymlikatrgvfpykgltrassrlhdelfrlilsp 976
QY 961 KFFDTPTGRILNRPFSKDMQEVDRALPFOAEMFIONVILVFFCGMAGVPMFLVAVGP 1020
DB 977 kffdtptgrilnrfpskdmdevdrvlpfqaemfionvllvfcgmlagvpmflvavgp 1036
QY 1021 LVILFVSLHIVSRVLELRLKRLDNITOSPFLSHITSSIOGLATTHAYNKGOEFLHROEL 1080
DB 1037 lvilfsvlhivsrvlrelkrldnitospflshitssiglatihaynkgeflhrygel 1096
QY 1081 LDNOQAPPELFTCAMRLAVLDLISALTGTGLMVLVHGQIPPAVAGLAISYAVOLT 1140
DB 1097 lddnqapflftcamravlldlislaltgtglmvlvmhgqipavaglaisyavolt 1156
QY 1141 GLFOFTVRLASETFAFTSVYRINHYIKTSLSEAPARIKKNAPSPDMPOEEVTFENAE 1200
DB 1157 glfotvrlasetfartsvyrinhyiktlslseaparikknapspdmpeevtfenaem 1216
QY 1201 RYRENPLVLKKSFTTKPKKIGIVGRTSGKSLGMAFLRVLELSGGCIKIDGVRI 1260
DB 1217 ryrenplvlkksfttkpkkigivrgtsgksslgmaflrvlelsggckidgvrisd 1276
QY 1261 IGLADRSKLSITPOEVLPGSVRSNLDPNQOTEPQOIMDALERTMKECIALDPLKLE 1320
DB 1277 igladrsklsitpoevlpgsvrsnldpnqotepqoimdalertmkecialdplkle 1336
QY 1321 SEVMENGDNFSGEBCROLCTARALRLRCKLILDEEATAMDETDLLIOETIRAFADCT 1380
DB 1337 sevmengdnfsgvrgqlctaralrlncklildeeatamdeddliigetirafadct 1396
QY 1381 MLTIAHRLHTVLSGDRIMVLAQGVVEFDPVSLSDSSRFYMPAAENKVAVK 1437
DB 1397 mltiahrlhtvlsgdrlmvlaqgvvefdpvsllsdssrfymfaaenkva 1453

RESULT 4

AA43542
ID AAY43542 standard; Protein: 1437 AA.
AC AAY43542;
XX 26-JAN-2000 (first entry)
DM
DE A human MPR-related ABC transporter designated MOAT-C.
XX
XX Human: MPR-related ABC transporter; MOAT protein; MOAT-C;
KM MOAT mediated transport; anticancer drug sensitivity;
KM transporter mediated cellular efflux; anticancer.
XX
XX Homo sapiens.
OS
XX MO9949735-A1.
FN
XX 07-OCT-1999.
PD
XX 26-MAR-1999; 99WO-US06644.
PE
XX 27-MAR-1998; 98US-0079759.
PR
PR 03-AUG-1998; 98US-0095153.
XX
XX (FOX-C-) FOX CHASE CANCER CENT.
PA
XX
XX Krush G, Lee K, Belinsky M, Bain L;
PI
XX WPI: 1999-610812/52.
DR
DR N-PSDB: AAZ30079.
XX
XX New transporter gene useful for screening for anti-cancer drugs -
PT
XX
XX Claim 16; Page 135-138; 153pp; English.
PS
XX
XX The present sequence represents a human MPR-related ABC transporter
CC (MOAT) protein, designated MOAT-C. The protein comprises a multi-domain
CC structure including a tandem repeat of nucleotide binding folds
CC appended C-terminal to a hydrophobic domain, having Walker A and B ATP
CC binding sites and several potential membrane spanning domains. The MOAT
CC nucleic acids are useful for screening a test compound for inhibition of
CC MOAT mediated transport, indicated by restoration of anticancer drug
CC sensitivity, which in turn causes a reduction of transporter mediated
CC cellular efflux of anticancer agents. MOAT DNA or RNA may be used as
CC probes to detect the presence or expression of genes encoding MOAT
CC proteins. Anti-MOAT antibodies are useful for detecting and quantitating
CC MOAT proteins.
CC
XX
SQ Sequence 1437 AA:

Query Match 99.8%; Score 7293; DB 20; Length 1437;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1434; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 241 alnytyrvlrgailltmafkklklnklsjgelincsndgrrmteaavgsllag 300
QY 301 PVVAILGMITVYVILIGPTGFIQSANVILFYPAAMFASRLTAIFRRKCYAADDERQKANE 360
Db 301 pvaillgmlyvillipgclfigsavfllfypammfasyrlayfrrckvaaderqkme 360
QY 361 VLTYSKFIKMYAWKAFQSOYOKIREERRIIEKAGYFOSTIVGAPVIVVYASVVFVS 420
Db 361 vltyskflkmyawvkaifsgvqkireeerlllekagyfqlgtvgyavplvvvyaavltsv 420
QY 421 HMTLGFDTLAAQAFVYVVFNSMPALVTPFSYKSLSEASVADRFKSLFLMEVHM 480
Db 421 hmtlgfdltaagafvrvvfnsmtflkvtpfsyvsleasvavdrfkslfimeevhmik 480
QY 481 NKPSAPHIKIEKKNAITLWDSHSSIQNSPKLTPKMKDKRAASRKKRKYOLQRTENQA 540
Db 481 nkpsaphikieknaitlwdshssiqnspkltpkmkdkraasrgkkekvrqlqrleqna 540
QY 541 VLABQKHLLDSDERPPEEEGKHHLGHRLQRTLHSDLEIOEKKVIGSGVSG 600
Db 541 vlaeqkghllldsderpseeegkhhlghrlqrtlshsdleieqgklyqicgsvsg 600
QY 601 KTSLSAILGQWTLLEGSAISGTFAYVAQAWINATLRDNIILRGKEDERYNSVNS 660
Db 601 ktslsailgqwtllegsaistgftayvaqawlnatlrdniilrgkederynsvlns 660
QY 661 CCLRPDLAILPSSDLTEIGERGANSQGORQRISTARALYSRSTIYILDDPISALDAHYG 720
Db 661 cclrpdlailpssdlteigeranlsqgrqrisaralysrstyilddpisaaldahyg 720
QY 721 NHIFNSAIRKHLSKSTVLFYHQLOLYVDCDEVIFPMKGCTERGTHERLMLNDVYATI 780
Db 721 nhifnsairkhlskstvlfyhqlolyvdcdevifpmkgctergtlherlmlndvyati 780
QY 781 FNNLLIGETPVEINSKRETSQSKSODKCPKTSIKKEKAVKEEGOLVLEEKSGCS 840
Db 781 fnnlligetpveinskretsqsksgksgdkpkiqsvkkekavkeegqlvgleeksgcs 840
QY 841 VPMVSYGYVYVIAAGGPLAFLVYIMALEFNLVNGSTABSTWMLSWIKSGSNTVTNENETS 900
Db 841 vpmvsygyvyvyaaggpflavlyimalenfnlvngstabstwmlswiksgsntvtrenets 900
QY 901 VSDSKNDPBNQYATYALSMANVMLIKATIGVYFVNGTLAASSRLDELFRRLRBSM 960
Db 901 vdsdkndpbnqyatyalsmanvmlikatigvyfvngtlaassrldelfrlrlbsm 960
QY 961 KFFDTPGRIINRFSKMDDEVRLPQAEFIONVILVFCVGMINGVPMFPAVAGP 1020
Db 961 kffdtptgrilnrfskmddevrllpqaeftionvilvfcvgmimgvpmfpavagp 1020
QY 1021 LVILFVSLIVSRVLIRELKRLDNTTOSPFSLHITSIOGLATIHAYNKGQEFLLRYOEL 1080
Db 1021 lvilfsvslivsrvlirelkrldnttospfslhittsioqlatihaynkqgefllryogel 1080
QY 1081 LDDNOAPPEFLPTCAMRWLAVALDLISALITTTGMLTYLMHGOIPRAYAGLAISTAVOLT 1140
Db 1081 lddnqapfflptcamrwlavrlidlisalitttgmltylmhgoiprayaglaistavolt 1140
QY 1141 GLFOFTVRLASTEARFVSVERINHYIKTSLSEAPARIKNKPSDPMPOEGEVTEPMEM 1200
Db 1141 glfoftvrlastearfvsverinhyiktlsleapariknkpsdpmppogevtetenem 1200
QY 1201 RYRENLPVLKKVSEFTIPKKEKIGIVGRSGKSSIGNALFRLVELSGGCIRIDGVRISD 1260
Db 1201 ryrenlpvlkkvseftipkkekigivgrsgkssignalfrlvelsggciridgvrisd 1260
QY 1261 IGLADRSKLSITIPQEPVLFSGTVSNLDPFNQYTEDQIMALERTHMKKECIAQPLKLE 1320
Db 1261 igladrsklsitipqepvlfsgtvsnldpfnqytedqimaletrthmkkeciaqplkle 1320
QY 1321 SEVMENGNFVSGEROLCIAIALRHCKIILDEATATAMDETQLIOETFREAFADCT 1380
Db 1321 sevmengnfvsgerolciaialrhckilildeatatamdetqlioetfreafadct 1380

Dh	387	icfvsgrmrlktaevigvnykcalvlnsacstsvgejvnlmsvdagrfmdlatylnmhs	446
Qy	300	gprvvallmgmtyvnlilgprgfglgsavfllfryammfzrlayrfrkrcvvaatdbrrvokm	359
Dh	447	aprlqvlla1y1wlnlgsprvsagvawmylmpvnavmamkcklvayahmkadarklkm	506
Qy	360	evltvylkfkimymwvafesqsovkireeberrllkragyfosittvgvavivvlnsvtfs	419
Dh	507	ellingklkvkljyawaefdkvlatrgeelkvllksaylsavgtftwctpflvalctfa	566
Qy	420	vhmftlged--lraoafvtyvtyfnpsntfplaktyprsvkslsbasvavdrfkslmevyn	477
Dh	567	vyvltidennlldeqtafivfalaflnllrflnllpwnsvsivagvskrlrlflfshlee	626
Qy	478	--mknkpap---hikiekmaatlamsdshssionspklrpkmkdkdrasrgkkekvrq	532
Dh	627	pdslerrpvkdgagcnsiclvnatfiv-----	653
Qy	533	lqtrhoavlaeogkghllldsderspseeegknhlghlrlqtrlshsldeloeklvq	592
Dh	654	-----arsdpp-----clnglftsfpegavla	675
Qy	593	icgsygsgrksllsalglomtllegslasgcfayvaqoawmlnatlndnllfckevdee	652
Dh	676	vvgvqvgqskslsallaeendkvegvalksgvayvvpqvgqndslrenllfqcgleep	735
Qy	653	rynsylvnscclrbdlalilpsrdltelgeegamlsqgororislaalxsdksrlyldbpl	712
Dh	736	yysvafgacallpdllelpsgrdtelgekvnlsgsqkqrvslarayvnadylfddpl	795
Qy	713	saldahgnhifnsal--rkhlkskvivlevthqoloyldcdevifpmksgcttergthel	770
Dh	796	savdahyghkhlfenvlygprkymknktrllvthsmyslpqvdlvlymsgqklsengsygel	855
Qy	771	mlnlgowdattf-----nnlll-----gempvvin	795
Dh	856	lardgataelrlryastegeqdeengvgyvgpgkkaekmenghnlvdsasgkqdyqls	915
Qy	796	skretsgsqoksdokdgprktsikrkeavkpreegolvleekgogsvpmsvtyvtoaaag	855
Dh	916	ssssygsdlsrhn-----staelqkaeakkeetwklmeadkaqtgvyklsywdymkaigl	972
Qy	856	plavltvlnafemlnvgstafstwmwlsywlkqsggnrtvrgnetvsvsmdnprmqvya	915
Dh	973	fistflsflf1fcmhvsalasnmwslwcdoplyngt---gehtkvrlsvygaljsgsl	1028
Qy	916	stvalsaavnwllkairgvvfyvgtlrassrlladelerrllrslpmkffdtptgtrilnr	975
Dh	1029	avfgymsavsi-----gglasrcihvdllhsilrpsmferrpssgnlvnrf	1076
Qy	976	skmdedevdvalpfoaemefionvlylpcvgmllagvprmwlvavglvllfsvlthsvyl	1035
Dh	1077	skeldvtdsmalpevklmfnslfntvngactvlllatpblaalllpprlgily---ffvqfy	1133
Qy	1036	l---relkrrldntospfshrtssioqatlatihaywnkgqeflhraroellddnovaflf	1092
Dh	1134	vaserqjkrleesvprpyshfnetllgyvslvrafceegrfllngsdklvdenqkayysi	1193
Qy	1093	camrwlaavrldlslalitttgmlvymhgolprpavaglaitsvaoqlglfofvyrlase	1152
Dh	1194	vamwlaavrleecygnclvfaalfavyshtslasaglvglsyagslyqvtlylwmvymase	1253
Qy	1153	tearfteverbthnykrlslsleapariknnapspdpqobevfennamkrenpnlvlyrk	1212
Dh	1254	mechnivalvelkeyset-ekaeppqldetacappsswqvqvavefnyrclylredlfdvtrh	1312
Qy	1213	vspftrkrekrlgtygrtsgsksslgmaelrvelsgcrlkrdgvrtsldgladrsklsl	1272
Dh	1313	invteggkvgvlygrtgagkslslglfrlfneseageglldglnlaklglnldlrftkl	1372
Qy	1273	tpogpvyfsgstvasnldpfnqoytieddoinalefthmkceliaqlpklasevmenndnsy	1332

D	b	1373	ipqpvlfsgslmmlidpfsgydeewerwlslelahnkdfvsaipdkldhecaegenlsv	1432
O	y	1333	GEROLICIAALLRHCKILILDETAAMTETDLLOETIREAFACCTMFIARHRTVL	1392
D	b	1433	gqrqlvclarrallrkrtkllvldeaavaadletddllqstlrftgcdctvltahrrintlm	1492
O	y	1393	GSDRIWLAOGAVEPDTPSVLLSNDSRYEAM	1425
D	b	1493	dylrvivldkgeidegyapsdll-qgrglfitysm	1524
RESULT 7				
I	D	AAW99894	standard; Protein; 1531 AA.	
X	X	AAW99894;		
A	C	10-JUN-1999	(first entry)	
D	T			
D	E		Human multidrug resistance-associated protein natural variant.	
K	M		Human; multidrug resistance-associated protein; MRP; cytotoxic drug;	
K	M		cancer; chemotherapy.	
X	X		Homo sapiens.	
O	S		US5891724-A.	
P	N		06-APR-1999.	
X	X		05-JUN-1995; 950S-0460907.	
P	F		05-JUN-1995; 950S-0460907.	
X	X		05-JUN-1995; 950S-0460907.	
P	R		27-OCT-1992; 920S-0966923.	
P	R		08-MAR-1993; 930S-0029340.	
P	R		26-OCT-1993; 930S-0141893.	
P	R		20-MAR-1995; 950S-0407207.	
X	P		(TOOH) UNIV QUEENS KINGSTON.	
X	P		Cole SPC, Deeley RG;	
X	P		WPI: 1999-253868/21.	
D	R		N-PSDB: AAXI9818.	
X	P		Protecting mammalian cells against cytotoxic drugs	
X	P		Claim 2; Column 79-86; 82pp; English.	
X	P		The present sequence represents a human multidrug resistance-associated	
C	C		protein (MRP). The present invention also describes a method for	
C	C		protecting a mammalian cell against the cytotoxicity of anthracyclines,	
C	C		epidopodophyllotoxins and Vinca alkaloids (A) by introducing into it a	
C	C		nucleic acid (I) that hybridizes under stringent conditions to a nucleia	
C	C		acid (II) that encodes an MRP protein (III). Introduction of (I)	
C	C		protects cells against cytotoxic effects of (A), particularly to protect	
C	C		normal cells with (I) can be used to screen for agents that affect	
C	C		transformed with (I) can be used to screen for agents that affect	
C	C		multidrug resistance or are directly toxic to multidrug resistant cells,	
C	C		i.e. potential therapeutics for multidrug-resistant cancers. Confering	
C	C		resistance to normal cells should allow an increase in the dose of (A)	
C	C		that can be administered safely.	
X	X		Sequence 1531 AA;	
S	O			
Q	y	100	PVDNAGLFGSCFTFWLSLARVAHKKGELSMEDYWSLKSHSSDVNCRRRLERLMOEBLE	159
D	b	209	pessaasfrilfwitgll-vryyrqplgsdwlswlnkedtsqgvpylvknmkkecak	267

```

OY 160 V-----GPDAA-----SLKRVATFCTRLL 181
Db 268 trkqpvkvvysskdpqkesskvaneveallvkspqkewmpslfvyktyfgyfym 327
OY 182 SIVCLMTIQLGSGSPAPMVKHLLFETQATESNLQYSLLVGLLFTFIVRMSLALMA 241
Db 328 sfffaiaidlmfmgpp-ikllikfyndtkapdqgyfyevllvtacqlvlyhygh 386
OY 242 LNYRTVRLRGALITMARCKILIKNIKERS--IGELINICSNDQORMEFAAVGSLLAG 299
Db 387 lcfvsgmrlkavlgavrykalvlnsarkstvgelivlnmsvdqgrlmdlatylnmlws 446
OY 300 GPVVAIILCMITNVILGTGTGLGSAVFILFYPAMMFASRLIYFRKCVAAATDEVQKKN 359
Db 447 aplqylalalyllwlnlpsvlagavmvmwlmvvnvnamkcktygvnamkckndrtiklmn 506
OY 360 EVLTATYIKFKYAWYKASQSVOKIREEERILEKAGYFOSITGVAVPIVYVIAVWPS 419
Db 507 eilngikvikiyawaekakvialrgeelkvikksaylsavglttwcctflvalcfta 566
OY 420 VHMFLGPD--LTAQAFTVTVVFNMSFPALKVTPSVKSLSEASVAVDRFKSLFMEVH 477
Db 567 vyvliidennlidaqtafvsalfoilfrlnlpmvissivqasvslkrtriflshoele 626
OY 478 --MKNKPAAP---HIKEMKNATIAMDSHSSIONSREKLFPKMKKDKRASGKREKVRQ 532
Db 627 psdlerfpvkdggsntslvratftw----- 653
OY 533 LQRTFHOAVLAQKCHLLDSDRPSPEEBGKHILHGLRLQRTLHSIDLEIOBGKLVG 592
Db 654 -----arsdpp-----tlngitfsipealva 675
OY 593 ICGSVSGKTSLSIALIQMTLLEGSIAISGTFAYVAOQAMILNATLRDNLFGKEYDEE 652
Db 676 vvgvgcgkssllsaliaemdkveghvalkgsavvypqawlnqdsirenlifgqleep 735
OY 653 RYNSVYNCCCLRPDLATIRPSSDLTFRIGERGANLSCGORISIALALSDRSITLIDPL 712
Db 736 yrrsvlgacallpdllelpsqdrteigekyvlsgqkqrvyslaayvnadilyfiddpl 795
OY 713 SALDAHVGNHIFNSAI--RKHLKSKTVFVTHQLOYLVDCCDEVIFMKEGCTERGTHEEL 770
Db 796 savdahvgkhlfenvigkgnkknktrilvchmsmylpqvadliymsgklsiemsgyel 855
OY 771 MNLNGDVATIR-----NNLL-----GERTPVEIN 795
Db 856 lardgafaeiflryasteqdaqeengvtvsgpgkaekqmengmvlvdsagkqlqrqls 915
OY 796 SKKETSSGOKSODKGRPTGSIKKKKAVKPREGOLVOLEKQGSVPMSVYVIOAAG 855
Db 916 ssssvsgdistrhn--staellqkaekaeetwklmeadkaqtgklsivwdynaklgl 972
OY 856 PLAFVIALFELNAGSTAFSTWMLSYWIKQSGNTVTYRGNETSVSDSKDNPHQYXA 915
Db 973 fislifslfmcnhsalaenyslwtddpivngt---gehtkvrslsvgalaisggl 1028
OY 916 SIYALSMVMLILKARGVVFVKGTLRASSRLHDELFRILRISPKKFFDPTPTGILNLF 975
Db 1029 avfgyssmavsl-----ggllasrchvdlhslrpsmfsferlpsgnlvmrf 1076
OY 976 SKDMDEVNRLPFOAEMFIQNVILVFCVGMTAGVPMFVLAVAGVLILFSLHIVSRVL 1035
Db 1077 skeldtvsamipevklmtgmslfnvlgacivlllatpilaailipplgily---ffvqrfly 1133
OY 1036 I---RELKRLDITQSPFLSHITSSIOGLATTHAYNKGCEFLHARQOELDDNQAFLEFT 1092
Db 1134 vassrgklrlevsvrpyshnelllysvvtrafeegerfihgdsikvdenqkaypsi 1193
OY 1093 CAMRLAVRLDLISALITTTGMLTVMHGOIPPAYAGLAISYAVOLGCLFOFTVRLASE 1152
Db 1194 vanrwlavrlaecvgncivlfaalfavlsrshisaglvgsavyslqvltlylmlvrmsee 1253

```

```

OY 1153 TEARTSEVERINHYIKTSLSEAPRIKKNAPSPDMPQEGEYFEENAKMYRENPLVLKK 1212
Db 1254 metnlvaverlkeyset-ekearpwqigetappswpvygrvefrnrclytriedlrfvllh 1312
OY 1213 VSFTIKPRKEIGIVGRTSGSKSLMGLFRLVELSGGCIRKIDGVRISPDIGLADLSKLSI 1272
Db 1313 lnvtlungekgvlgvrtgagkscltglfrinesaegeildglnlaklghlndrlfkltl 1372
OY 1273 IPOEPVFSGTVRSNLDPPNOYTEDQIWDALERTHMKECIAOLPLKLESEVMEGDNFSV 1332
Db 1373 lpgdbylftsgslrmldpfsfsgsdeevwtstelahktdfvsalpkldhecegegensv 1432
OY 1333 GEROLCLTARALLRCKIILIDEAATAMDTERTDILQETIRAFADCMULTAHHLHPLV 1392
Db 1433 gqrqlvclarrallrktklilvdeataavldetdillqslrtfqcdevctvlahrlntlm 1492
OY 1393 GSDRIWVLAQGVFEFDPSPVLLSDSSRFYAM 1425
Db 1493 dylrvlvdkgelgeigpapsdl1-qgrqlfysm 1524

RESULT 8
AAW74471
ID AAW74471 standard; Proteins: 1531 AA.
XX
AC AAW74471;
Df 18-MAY-1999 (first entry)
XX
DE Human multidrug resistance-associated protein variant.
KM Multidrug resistance-associated protein; MDR; human; diagnosis;
KM MDR tumour cell identification; cancer therapy.
OS Homo sapiens.
XX
PN US5882875-A.
XX
PD 16-MAR-1999.
XX
PF 05-JUN-1995; 95US-0462109.
XX
PR 05-JUN-1995; 95US-0462109.
PR 27-OCT-1992; 92US-0966923.
PR 08-MAR-1993; 93US-0029340.
PR 26-OCT-1993; 93US-0141893.
PR 20-MAR-1995; 95US-0407207.
XX
PA (TOOH ) UNIV QUEBENS KINGSTON.
XX
PI Cole SPC; Deeley RG;
XX
DR WPI; 1999-214061/18.
XX
DR N-PSDB; AAX21977.
XX
PT Identifying a multidrug resistant tumour cell by contacting the cell
PT with an antibody/antigen-binding fragment - which binds to an
PT expressed protein encoded by multidrug resistance-associated protein
PT (MRP) nucleic acid
XX
PS Claim 3; Column 69-80; 80pp; English.
XX
CC This sequence is the human multidrug resistance-associated (MDR)
CC protein. The invention relates to a method for identifying a multidrug
CC resistant (MDR) tumour cell. Compositions and methods utilising the MDR
CC proteins can be used to treat patients with tumours displaying multidrug
CC resistance, particularly those displaying resistance to antineoplastic
CC epipodophylotoxins, vinca alkaloids, and hydrophobic drugs. The methods
CC for inhibiting/killing a MDR tumour cell can be useful for treating
CC breast cancer, leukemias, fibrosarcomas, cervical cancer, gliomas,
CC thymas, neuroblastomas and lung cancer. The MDR DNA sequences when
CC labeled are useful as molecular probes for diagnosing multidrug
CC resistance of a tumour (using cells from a tumour biopsy) and for

```


XX 01-APR-1998 (first entry)
XX Human canalicular multispecific organic anion transporter protein.
XX
XX
XX Canalicular multispecific organic anion transporter protein;
XX cMOAT protein; ATP-binding cassette transporter family; ABC transporter;
XX hepatobiliary excretion; multidrug resistance-associated protein;
XX cMOAT protein activity; multidrug resistance-related protein; MDR-1;
XX Dubin-Johnson disease; Rotor disease.
XX
XX Homo sapiens.
XX
XX
XX Key
XX Modified-site 7..9 Location/Qualifiers
XX /note= "Asn is predicted to be N-glycosylated"
XX 26..49
XX /note= "putative transmembrane region"
XX 74..87
XX /note= "putative transmembrane region"
XX 106..118
XX /note= "putative transmembrane region"
XX 130..148
XX /note= "putative transmembrane region"
XX 166..184
XX /note= "putative transmembrane region"
XX 319..341
XX /note= "putative transmembrane region"
XX 361..379
XX /note= "putative transmembrane region"
XX 440..456
XX /note= "putative transmembrane region"
XX 463..480
XX /note= "putative transmembrane region"
XX 546..566
XX /note= "putative transmembrane region"
XX 583..605
XX /note= "putative transmembrane region"
XX 671..679
XX /note= "Walker A, B or signature sequence"
XX 761..771
XX /note= "Walker A, B or signature sequence"
XX 781..786
XX /note= "Walker A, B or signature sequence"
XX 976..996
XX /note= "putative transmembrane region"
XX 911..913
XX /note= "Asn is predicted to be N-glycosylated"
XX 1025..1050
XX /note= "putative transmembrane domain"
XX 1101..1120
XX /note= "putative transmembrane region"
XX 1022..1039
XX /note= "putative transmembrane region"
XX 1212..1232
XX /note= "putative transmembrane region"
XX 1238..1251
XX /note= "putative transmembrane region"
XX 1334..1342
XX /note= "Walker A, B or signature sequence"
XX 1437..1447
XX /note= "Walker A, B or signature sequence"
XX 1457..1462
XX /note= "Walker A, B or signature sequence"
XX
XX W09731111-A2.
XX
XX PD 28-AUG-1997.
XX
XX 21-FEB-1997; 97WO-NL00079.
XX
XX 22-FEB-1996; 96EP-0200460.
XX

PA (MEDI-) ACAD MEDISCH CENT AMSTERDAM.
PA (HEEN-) HET NEDERLANDS KANKER INST.
PA (INTR-) INTRIGENE BV.
XX
XX Borst P, Bosma PJ, Evers R, Oude Elferink RPJ;
PI Paulusma CC;
PI
XX WPI: 1997-435163/40.
DR N-PSDB: AAT94023.
XX
XX DNA encoding human and rat canalicular multispecific organic anion
PT transporter proteins - useful for diagnosis and treatment of
PT Dubin-Johnson disease and Rotor disease
XX
XX
PS Disclosure: Fig 10; 106pp; English.
XX
XX The present sequence represents a novel canalicular multispecific
CC organic anion transporter (cMOAT) protein. The protein is a new member of
CC the ATP-binding cassette (ABC) transporter family. The ATP dependent
CC cMOAT transporter system mediates hepatobiliary excretion in the liver.
CC cMOAT may be a liver-specific homologue of multidrug
CC resistance-associated protein. The nucleic acids are used to provide
CC cells with cMOAT protein activity. cMOAT protein activity in cells can be
CC enhanced by increasing the level of glutathione, glucuronide and/or
CC sulphate. Antisense constructs, especially derived from another multidrug
CC resistance (MDR)-related protein, e.g. MDR-1, to the nucleic acids and
CC vectors can be used to decrease the level of cMOAT in a cell. The nucleic
CC acids and proteins can be used especially in diagnosis of Dubin-Johnson
CC disease, Rotor disease or another disease involving cMOAT. The cMOAT gene
CC may also be used as a selectable marker gene.
XX
SQ Sequence 1545 AA:

Query Match 28.5%; Score 2082; DB 18; Length 1545;
Best Local Similarity 34.5%; Pred. No. 1.2e-165;
Matches 502; Conservative 246; Mismatches 488; Indels 218; Gaps 27;

QY 99 HPVDNAGLFSCMTFSWLSLARVAHKGELSMEDVWSLSKHESSDVNCRRLRMOEIN 158
DB 193 npsiasfssityswysillkykr-pltledvewdeemkktlvskfethmkrelq 251
QY 159 EV-----GPDASLRV-----VW-- 172
DB 252 karalqrrqekssqngsgarlpjlnknsgsqdalvledvekkksgctkdvpxswlm 311
QY 173 --IFCRTLIL--SIVCLMITQLAGSGPAFWKHLLEYQTATSNLQYSLILVGLILT 228
DB 312 kalfttfymvllksfllklniditfvsbpq-llkllisfasdrdyvwgylcaillfta 370
QY 229 EIVRSWSLALFWALNYRTGVRLGAILTMAFKILKLNI--KEKSLGELINICSDGOR 286
DB 371 aligfcqlqcyfqlcfklkvrtaimasvykaltlnlarkeytvgectvlnmsvdaqk 430
QY 287 MFEAAVAGSLAGPYNVALGMIVYVILGPTGLGSNAVFIFRYAMMPASLTATYFRK 346
DB 431 lmdvnlfmhmlwssvlgjvlisflwrelgsvlgvymvltvldinalstksxtlqv 490
QY 347 CVAATDERVOKNEVLTYIKFKIMYAWYKAFSQSVOKTIREERRILKAGYPOSTIVGA 406
DB 491 mknkdkrtiklmellsgkiklkyfawepstfdqynlkkkalknllafsqdqcvlftf 550
QY 407 PIVVIVASVTFVSMTLGF--LTAAQFTVTVFVNSMTFALKYTPSVKSLASAVAV 464
DB 551 qltpvlsvvtfsvyvlvdsnlldaqkafstltlfnlrfplmnpmmismlgasvat 610
QY 465 DRFKSLFMEEVYHMKNNRPASPH-----IKTEKNATILAMDSHSSIONSPLTRKMKKD 519
DB 611 erlekylgddl-----dtsalrhdcnfcfkamqfseastf----- 646
QY 520 KRASGRKEKYVLOLQTEHOAVLABOKGHLILDSDRSPSPREEEGKHILGHLRLQRTLH 579
DB 647 -----ehds-----catvr 655


```

Db 242 gpkessvdaaneaveallvkspqkewmps1fkvlykftgpyflmsffekajhndlmmsfsgp 301
QY 198 AFMWHLELYQATSNIOYSLILVLGLLFTIVRSKSLATWALNRYRGVLRGAILTM 257
Db 302 q-llkllikftvdapdqwyfyfvl1lftvacqltlvhyfhhcftvsgmktkav1ga 360
QY 258 AFKIKIKNIKIKERS--LGEELINICNDGQRMFEAAVGSLLAGPVAALMINTVIL 315
Db 361 vyrkalvltnsrksstvgelvnmsvdqgrfmdlaelyimmsaprlqyl1l1w1n1 420
QY 316 GPTEGLSAVFLLEYFAMAFASRLATYFRRCVAAATDERVOKMNEVLTYIKPKMYAWK 375
Db 421 gpsv1agayvnm1mrvpnaamktktyvahmkskdnr1k1mnel1ng1kv1k1yawe1 480
QY 376 AFSSQVOKIREEERILKACGFOSITVGAPIVVIVIASVYTFESHM1LGEF--LTAQA 433
Db 481 arkdav1arqeel1kksaylsavgtftwvcrpf1valctfayvltldem1l1daqta 540
QY 434 FTVVTFNSMTFALKVTPFSYKSLSEASVAVDRFSLFMEEVH--MIKNKPASP---HI 488
Db 541 fvs1alf1l1rpl1l1pmv1ss1vgsavslkr1l1flshel1epds1er1rpkdgsg1n 600
QY 489 KIEMKATLAMDSSHS1IONSFKLTPKMKKDKRASRCKEKYRQLRTEHQA1LAEQKH 548
Db 601 sltvna1ftw----- 611
QY 549 LLLDDEPSPPEEGKHHHGLHLALQRTLHSDLEIOGK1VYG1CSGSGKTSLSAI 608
Db 612 -----arsdp-----tlng1lts1pegalvavavqvcgk1sl1s1l 649
QY 609 LGQMTLEGS1AISCTFAYVAQAM1LNTATLND1LFCGEYDEERYNSV1LNCCLRPDLA 668
Db 650 laemdkveg1hva1kgs1vayvpgaw1qnds1ren1l1f1gcq1leep1ry1sv1dgc1alp1de 709
QY 669 ILPSDLTEIGERGANLSGGQR1SLARALYSDBRS1Y1LDDP1SALDAHVGNH1FNSAI 728
Db 710 l1psg1rte1qeky1n1sg1qk1r1s1ar1ay1s1n1d1y1l1f1d1p1s1av1d1ah1vk1h1fenv1 769
QY 729 --RKHLKSKTYLFVTHO1Y1VDODEV1FMKGC1TERTGTH1ELM1NLNDVAT1F----- 781
Db 770 gpkgm1k1k1r1l1v1h1sm1sy1lpqvd1l1vm1sg1k1sem1sy1gell1ard1q1a1ef1r1tyas 829
QY 782 -----NNLL-----G1ETP1V1E1N1SK1K1ETS1G1SQ1KS1OD1K 811
Db 830 teqeg1dae1ng1ty1sg1qk1eak1qem1ng1l1v1ds1ag1k1q1r1q1l1ss1ss1sg1s1r1h1n1-- 887
QY 812 PKTGS1K1KEK1AVK1PEEGOLV1LEEK1G1SG1VPW1SYGV1Y1OAA1G1PLAT1V1MAL1PLN1VG 871
Db 888 -sta1el1q1aa1eak1eet1wk1lmead1q1t1g1vk1s1vy1wd1ym1ka1q1l1f1s1f1s1f1-1f1m1cn1h1v 945
QY 872 STAFSTM1LWSWIK1GSGNT1YTRNG1NET1SVSD1MKDN1PM1OY1YAS1IYAL1MSM1VML1LKA1 931
Db 946 sala1ny1w1sl1v1t1cd1p1v1ng1t1---qeh1k1vr1s1y1ga1l1sg1l1av1f1g1ym1sv1s1t1----- 997
QY 932 RGQVVEK1TLRAS1RLHDEL1FR1RLSPMK1FED1TPTG1RLN1RFS1KDM1EDV1R1PFOAE 991
Db 998 -----g1l1as1r1c1h1nd1l1hs1l1r1ps1m1f1er1cp1sg1n1l1vr1f1sk1eld1v1sm1p1ev1k 1050
QY 992 MFIQW1L1V1F1C1VGM1AG1VFPW1L1VAVGP1VL1FSV1LH1VSHV1L1--RELK1RLDN1TQ1S 1048
Db 1051 m1fmg1sl1f1nv1g1ac1yl1l1at1p1aa1l1pp1l1g1ly1---ffv1q1f1y1ass1rq1l1r1les1v1s1 1107
QY 1049 PELSH1TS1IG1LAT1IH1V1NK1GOE1LH1RY1QEL1LDN1OAP1F1ET1CAM1RL1AV1RLDL1S1A 1108
Db 1108 p1v1s1h1f1n1e1ll1gv1s1l1r1a1f1e1e1g1er1f1h1gs1d1k1v1den1q1k1ay1ps1v1aan1w1l1av1l1ec1v1nc 1167
QY 1109 L1TT1G1L1M1V1H1G1Q1P1AY1AG1A1S1VAVOL1TGL1FO1F1V1R1L1ASE1F1EAP1T1S1VER1N1H1Y1K 1168
Db 1168 l1v1l1fa1l1av1s1r1h1s1gl1v1g1s1ys1l1q1v1t1y1n1v1l1y1ms1sem1et1n1v1a1er1l1key1se 1227
QY 1169 T1LSE1AP1R1K1K1AP1SP1MP1OGE1V1TF1ENAE1MR1RENT1PLV1LK1V1S1F1IK1PEK1IG1VGR 1228
Db 1228 t-ek1eap1q1d1q1et1r1p1ss1w1p1q1v1g1r1ve1f1r1nc1l1y1r1ed1l1d1v1l1r1h1n1v1l1nge1k1v1g1vr 1286

```

```

QY 1229 TGSGKS1IGMALFRLVEL1SGC1K1ID1GV1R1SD1GLAD1RSK1S1I1POEP1V1FS1G1V1RS1NL 1288
Db 1287 t1ga1gk1s1l1g1l1f1r1l1nea1e1g1l1d1g1n1l1ak1g1h1d1r1fk1t1l1p1q1d1v1l1sg1s1l1m1l 1346
QY 1289 DPFNQ1TDEQ1WDALERT1HMK1E1A1Q1LPLK1ESE1V1M1ENG1DN1F1SG1EROL1CTAR1ALL1RHC 1348
Db 1347 d1f1sg1s1de1ev1ts1el1ah1k1f1sa1l1p1k1d1h1e1ca1eg1en1s1vg1r1q1v1cl1ar1l1r1kt 1406
QY 1349 K1L1L1DE1TAM1D1T1DL1L1Q1ET1RE1AP1D1CT1M1T1A1R1L1HT1V1L1G1SD1R1M1V1LA1Q1VVE1F 1408
Db 1407 k1l1v1dea1ta1vad1l1et1dd1l1g1st1l1rt1q1ed1cv1l1a1h1n1l1m1d1y1tr1v1l1d1k1e1g1ey 1466
QY 1409 D1P1SV1L1S1ND1SS1RFYAM 1425
Db 1467 g1ap1s1d1l1-q1r1gl1f1ysm 1482

RESULT 13
AAR54928
ID AAR54928 standard; Protein; 1531 AA.
XX
AC AAR54928;
XX
XX 14-OCT-1994 (first entry)
DT
XX
DE Multidrug resistance protein.
XX
XX Multidrug resistance protein; MRP; H69AR; cancer cell line;
KM stem cell; cardiac muscle; transgenic animal.
XX
OS Homo sapiens.
XX
XX WO9410303-A.
XX
XX 11-MAY-1994.
XX
XX 27-OCT-1993; 93MO-CA00439.
XX
XX 27-OCT-1992; 92US-0966923.
XX
XX 08-MAR-1993; 93US-0029340.
XX
XX (TOOH ) UNIV QUEBENS KINGSTON.
XX
XX Cole SPC, Deeley RG;
XX
XX WPI: 1994-167460/20.
XX
XX DR N-PSDB; AAO65377.
XX
XX Multi-drug resistance gene - encodes protein capable of
PT conferring multi-drug resistance on cells, useful in diagnostic
PI and treatment methods
XX
XX PS Disclosure; Page 69-74; 101pp; English.
XX
XX The multidrug resistant cancer cell line H69AR (ATCC CRL 11350) was
CC used to identify cDNA encoding a novel protein associated with
CC multidrug resistance, MRP. MRP may be expressed in e.g.
CC hematopoietic stem cells or cardiac muscle, or in
CC transgenic animals, or can be used to raise antibodies.
XX
XX SO Sequence 1531 AA;

Query Match 28.4%; Score 2075; DB 15; Length 1531;
Best Local Similarity 34.5%; Pred. No. 4.5e-165;
Matches 487; Conservative 261; Mismatches 481; Indels 184; Gaps 21.

QY 100 PYDNA1G1SCMT1FWS1L1SLAR1A1H1K1G1EL1SM1ED1V1SL1K1H1ES1DV1N1CR1L1ER1L1Q1E1L1NE 159
Db 209 p1ess1as1f1s1r1l1f1w1l1l1-v1ry1r1p1l1eg1s1d1w1l1k1ed1s1eq1v1v1p1v1k1m1w1k1e1c1a1 267
QY 160 V-----GPDA-----SLRRVW1F1C1R1L1L 181

```


Tue Nov 13 11:55:00 2001

us-09-528-031-2.rag

Page 19

```

      |::||::|: || || ||::|
Db 1493 dytrvivldkgeiqeygapsdl-qqrqlfysm 1524

```

Search completed: November 9, 2001, 17:45:56
Job time: 6854 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:53:52 ; Search time 34.17 Seconds
(without alignments)
946.363 Million cell updates/sec

Title: US-09-528-031-2

Perfect score: 7308

Sequence: 1 MKDIDGKKEIIIPSGYRSV.....DSSRFYAMFAAEKNAVKG 1437

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7308	100.0	1437	3	US-09-061-400-2
2	7298	99.9	1453	2	US-09-001-273-2
3	7298	99.9	1453	4	US-08-843-459A-2
4	2082	28.5	1531	1	US-08-463-092B-4
5	2082	28.5	1531	2	US-08-462-109A-4
6	2082	28.5	1531	2	US-08-460-907B-4
7	2082	28.5	1531	3	US-08-463-179A-4
8	2082	28.5	1531	3	US-08-461-384B-4
9	2075	28.4	1531	1	US-08-141-893-2
10	2075	28.4	1531	1	US-08-463-092B-2
11	2075	28.4	1531	2	US-08-462-109A-2
12	2075	28.4	1531	2	US-08-460-907B-2
13	2075	28.4	1531	3	US-08-463-179A-2
14	2075	28.4	1531	3	US-08-461-384B-2
15	2075	28.4	1531	3	US-08-407-207A-2
16	2055.5	28.1	1528	1	US-08-463-092B-6
17	2055.5	28.1	1528	2	US-08-462-109A-6
18	2055.5	28.1	1528	2	US-08-460-907B-6
19	2055.5	28.1	1528	3	US-08-463-179A-6
20	2055.5	28.1	1528	3	US-08-461-384B-6
21	2013.5	27.6	1622	4	US-08-972-927-6
22	1991	27.2	1621	4	US-08-972-927-3
23	1759	24.1	1581	4	US-08-726-320-3
24	1759	24.1	1581	4	US-09-208-716-3
25	1757.5	24.0	1580	4	US-08-726-320-1
26	1757.5	24.0	1580	4	US-09-208-716-1
27	1732	23.7	1581	4	US-08-726-320-4

28	1732	23.7	1581	4	US-09-208-716-4	Sequence 4, Appl.
29	1730	23.7	1477	3	US-08-492-459-10	Sequence 10, Appl
30	1730	23.7	1477	3	US-08-423-752-10	Sequence 10, Appl
31	1730	23.7	1477	3	US-08-945-994-3	Sequence 3, Appl
32	1730	23.7	1477	4	US-08-716-873-24	Sequence 24, Appl
33	1730	23.7	1477	4	US-09-368-431-24	Sequence 24, Appl
34	1669.5	22.8	1581	2	US-08-404-531B-6	Sequence 6, Appl
35	1669.5	22.8	1581	3	US-08-476-900A-6	Sequence 6, Appl
36	1669.5	22.8	1581	3	US-08-488-546A-6	Sequence 6, Appl
37	1669.5	22.8	1582	2	US-08-404-531B-9	Sequence 9, Appl
38	1669.5	22.8	1582	3	US-08-476-900A-9	Sequence 9, Appl
39	1669.5	22.8	1582	3	US-08-488-546A-9	Sequence 9, Appl
40	1669.5	22.8	1582	4	US-08-726-320-5	Sequence 5, Appl
41	1669.5	22.8	1582	4	US-09-208-716-5	Sequence 5, Appl
42	1528.5	20.9	1498	2	US-08-404-531B-28	Sequence 28, Appl
43	1528.5	20.9	1498	3	US-08-476-900A-28	Sequence 28, Appl
44	1528.5	20.9	1498	3	US-08-488-546A-28	Sequence 28, Appl
45	1513	20.7	1548	1	US-08-463-092B-7	Sequence 7, Appl

ALIGNMENTS

```
RESULT
US-09-061-400-2
; Sequence 2, Application US/09061400
; Patent No. 6077936
; GENERAL INFORMATION:
; APPLICANT: SHYJAN, Andrew
; TITLE OF INVENTION: NOVEL MULTIDRUG RESISTANCE-ASSOCIATED
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/061,400
; FILING DATE: 16-APRIL-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth A. Hanley
; REGISTRATION NUMBER: 33,505
; REFERENCE/DOCKET NUMBER: MNI-056CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-061-400-2

Query Match      100.0%; Score 7308; DB 3; Length 1437;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MDIDIGKEIIPSGYRSVREKTSSTGTHRDSDSKFRTRPLECODALETAARAGIS 60
      |||||||
DB      1 MDIDIGKEIIPSGYRSVREKTSSTGTHRDSDSKFRTRPLECODALETAARAGIS 60
QY      61 LDASMSQRIIDDEHPKCKYHNGLSALKPIRTCTCKHNPVDNAGLFSGMTSTWLSLAR 120
```

D	b	61	LDA\$M\$Q\$RLDLEB\$H\$K\$G\$K\$Y\$H\$H\$B\$S\$A\$K\$I\$R\$T\$T\$C\$K\$H\$Q\$H\$P\$V\$D\$A\$G\$F\$S\$C\$M\$T\$P\$S\$W\$S\$A\$R	120
Q	y	121	VA\$K\$G\$E\$S\$M\$E\$D\$V\$S\$T\$S\$K\$H\$E\$S\$S\$D\$V\$N\$C\$R\$R\$E\$R\$L\$M\$Q\$E\$L\$N\$E\$V\$P\$D\$A\$S\$L\$R\$V\$W\$J\$E\$C\$R\$T\$R\$L	180
D	b	121	VA\$K\$G\$E\$S\$M\$E\$D\$V\$S\$T\$S\$K\$H\$E\$S\$S\$D\$V\$N\$C\$R\$R\$E\$R\$L\$M\$Q\$E\$L\$N\$E\$V\$P\$D\$A\$S\$L\$R\$V\$W\$J\$E\$C\$R\$T\$R\$L	180
Q	y	181	L\$T\$Y\$C\$L\$M\$T\$O\$A\$G\$S\$G\$P\$A\$W\$K\$H\$E\$L\$E\$T\$O\$A\$T\$E\$S\$N\$I\$O\$Y\$S\$T\$U\$Y\$G\$L\$T\$E\$T\$E\$V\$R\$M\$S\$A\$L\$W	240
D	b	181	L\$T\$Y\$C\$L\$M\$T\$O\$A\$G\$S\$G\$P\$A\$W\$K\$H\$E\$L\$E\$T\$O\$A\$T\$E\$S\$N\$I\$O\$Y\$S\$T\$U\$Y\$G\$L\$T\$E\$T\$E\$V\$R\$M\$S\$A\$L\$W	240
Q	y	241	A\$L\$N\$R\$T\$G\$R\$U\$E\$G\$A\$L\$L\$T\$P\$A\$F\$K\$K\$I\$T\$K\$N\$K\$S\$G\$E\$L\$I\$N\$I\$C\$S\$N\$D\$Q\$M\$F\$E\$A\$A\$A\$G\$S\$T\$A\$G\$	300
D	b	241	A\$L\$N\$R\$T\$G\$R\$U\$E\$G\$A\$L\$L\$T\$P\$A\$F\$K\$K\$I\$T\$K\$N\$K\$S\$G\$E\$L\$I\$N\$I\$C\$S\$N\$D\$Q\$M\$F\$E\$A\$A\$A\$G\$S\$T\$A\$G\$	300
Q	y	301	P\$V\$A\$N\$I\$G\$M\$T\$V\$V\$T\$L\$G\$T\$G\$E\$L\$G\$S\$A\$V\$F\$L\$T\$P\$A\$M\$F\$S\$R\$L\$T\$A\$T\$R\$K\$C\$V\$A\$T\$D\$E\$R\$V\$O\$K\$N\$E	360
D	b	301	P\$V\$A\$N\$I\$G\$M\$T\$V\$V\$T\$L\$G\$T\$G\$E\$L\$G\$S\$A\$V\$F\$L\$T\$P\$A\$M\$F\$S\$R\$L\$T\$A\$T\$R\$K\$C\$V\$A\$T\$D\$E\$R\$V\$O\$K\$N\$E	360
Q	y	361	V\$L\$T\$Y\$K\$F\$K\$M\$A\$W\$K\$A\$P\$S\$O\$S\$O\$K\$R\$E\$E\$R\$R\$I\$E\$K\$A\$G\$F\$O\$S\$I\$T\$G\$V\$A\$P\$I\$V\$V\$A\$S\$V\$T\$F\$V	420
D	b	361	V\$L\$T\$Y\$K\$F\$K\$M\$A\$W\$K\$A\$P\$S\$O\$S\$O\$K\$R\$E\$E\$R\$R\$I\$E\$K\$A\$G\$F\$O\$S\$I\$T\$G\$V\$A\$P\$I\$V\$V\$A\$S\$V\$T\$F\$V	420
Q	y	421	H\$M\$T\$G\$F\$D\$L\$T\$A\$A\$O\$F\$T\$V\$V\$T\$V\$F\$N\$S\$M\$T\$F\$A\$L\$K\$V\$T\$P\$S\$V\$K\$S\$E\$A\$S\$V\$A\$D\$R\$K\$S\$L\$T\$M\$E\$E\$V\$H\$M\$K	480
D	b	421	H\$M\$T\$G\$F\$D\$L\$T\$A\$A\$O\$F\$T\$V\$V\$T\$V\$F\$N\$S\$M\$T\$F\$A\$L\$K\$V\$T\$P\$S\$V\$K\$S\$E\$A\$S\$V\$A\$D\$R\$K\$S\$L\$T\$M\$E\$E\$V\$H\$M\$K	480
Q	y	481	N\$K\$P\$A\$P\$H\$K\$I\$E\$M\$K\$A\$T\$A\$M\$O\$S\$H\$S\$I\$O\$N\$S\$K\$R\$L\$P\$K\$M\$K\$D\$A\$R\$G\$K\$E\$K\$Y\$O\$L\$O\$R\$T\$E\$H\$O\$A	540
D	b	481	N\$K\$P\$A\$P\$H\$K\$I\$E\$M\$K\$A\$T\$A\$M\$O\$S\$H\$S\$I\$O\$N\$S\$K\$R\$L\$P\$K\$M\$K\$D\$A\$R\$G\$K\$E\$K\$Y\$O\$L\$O\$R\$T\$E\$H\$O\$A	540
Q	y	541	V\$L\$A\$Q\$K\$G\$H\$L\$D\$D\$E\$R\$P\$S\$E\$E\$B\$E\$K\$H\$H\$G\$H\$L\$R\$L\$Q\$R\$T\$H\$S\$I\$D\$E\$L\$O\$E\$K\$V\$G\$S\$V\$G\$G	600
D	b	541	V\$L\$A\$Q\$K\$G\$H\$L\$D\$D\$E\$R\$P\$S\$E\$E\$B\$E\$K\$H\$H\$G\$H\$L\$R\$L\$Q\$R\$T\$H\$S\$I\$D\$E\$L\$O\$E\$K\$V\$G\$S\$V\$G\$G	600
Q	y	601	K\$T\$S\$L\$I\$A\$L\$O\$M\$T\$L\$E\$B\$S\$I\$A\$S\$G\$F\$E\$A\$V\$A\$Q\$A\$M\$T\$A\$N\$T\$L\$D\$N\$T\$L\$F\$K\$E\$V\$D\$E\$E\$R\$N\$S\$V\$L\$M\$S	660
D	b	601	K\$T\$S\$L\$I\$A\$L\$O\$M\$T\$L\$E\$B\$S\$I\$A\$S\$G\$F\$E\$A\$V\$A\$Q\$A\$M\$T\$A\$N\$T\$L\$D\$N\$T\$L\$F\$K\$E\$V\$D\$E\$E\$R\$N\$S\$V\$L\$M\$S	660
Q	y	661	C\$C\$L\$R\$P\$D\$L\$I\$P\$S\$D\$L\$T\$E\$I\$G\$E\$R\$G\$A\$N\$S\$G\$G\$O\$R\$O\$R\$I\$S\$L\$A\$R\$A\$Y\$S\$D\$R\$S\$I\$Y\$L\$D\$D\$P\$L\$S\$A\$L\$D\$A\$H\$V\$G	720
D	b	661	C\$C\$L\$R\$P\$D\$L\$I\$P\$S\$D\$L\$T\$E\$I\$G\$E\$R\$G\$A\$N\$S\$G\$G\$O\$R\$O\$R\$I\$S\$L\$A\$R\$A\$Y\$S\$D\$R\$S\$I\$Y\$L\$D\$D\$P\$L\$S\$A\$L\$D\$A\$H\$V\$G	720
Q	y	721	N\$H\$I\$N\$S\$A\$I\$R\$K\$H\$L\$K\$K\$Y\$V\$L\$P\$Y\$H\$O\$Q\$Y\$V\$D\$D\$E\$V\$Y\$P\$A\$K\$E\$G\$I\$T\$E\$R\$G\$H\$E\$L\$M\$N\$G\$Y\$A\$T	780
D	b	721	N\$H\$I\$N\$S\$A\$I\$R\$K\$H\$L\$K\$K\$Y\$V\$L\$P\$Y\$H\$O\$Q\$Y\$V\$D\$D\$E\$V\$Y\$P\$A\$K\$E\$G\$I\$T\$E\$R\$G\$H\$E\$L\$M\$N\$G\$Y\$A\$T	780
Q	y	781	F\$N\$N\$T\$L\$G\$T\$P\$E\$V\$E\$N\$S\$K\$K\$E\$S\$G\$S\$O\$K\$S\$O\$D\$K\$P\$K\$T\$G\$S\$I\$K\$K\$E\$A\$V\$P\$E\$G\$O\$L\$V\$O\$L\$E\$E\$K\$G\$O\$S	840
D	b	781	F\$N\$N\$T\$L\$G\$T\$P\$E\$V\$E\$N\$S\$K\$K\$E\$S\$G\$S\$O\$K\$S\$O\$D\$K\$P\$K\$T\$G\$S\$I\$K\$K\$E\$A\$V\$P\$E\$G\$O\$L\$V\$O\$L\$E\$E\$K\$G\$O\$S	840
Q	y	841	V\$P\$W\$S\$Y\$G\$Y\$I\$O\$A\$A\$G\$P\$L\$A\$F\$L\$V\$M\$A\$L\$F\$M\$L\$N\$G\$S\$T\$A\$F\$S\$T\$W\$M\$L\$S\$Y\$M\$I\$K\$O\$G\$S\$G\$N\$T\$Y\$T\$R\$G\$E\$N\$S	900
D	b	841	V\$P\$W\$S\$Y\$G\$Y\$I\$O\$A\$A\$G\$P\$L\$A\$F\$L\$V\$M\$A\$L\$F\$M\$L\$N\$G\$S\$T\$A\$F\$S\$T\$W\$M\$L\$S\$Y\$M\$I\$K\$O\$G\$S\$G\$N\$T\$Y\$T\$R\$G\$E\$N\$S	900
Q	y	901	V\$D\$S\$M\$K\$O\$N\$P\$H\$O\$Y\$A\$S\$Y\$A\$T\$S\$M\$A\$M\$L\$T\$K\$R\$G\$V\$P\$V\$K\$G\$L\$R\$A\$S\$R\$R\$H\$E\$L\$P\$R\$T\$R\$S\$M	960
D	b	901	V\$D\$S\$M\$K\$O\$N\$P\$H\$O\$Y\$A\$S\$Y\$A\$T\$S\$M\$A\$M\$L\$T\$K\$R\$G\$V\$P\$V\$K\$G\$L\$R\$A\$S\$R\$R\$H\$E\$L\$P\$R\$T\$R\$S\$M	960
Q	y	961	K\$E\$F\$D\$T\$P\$T\$G\$R\$L\$N\$F\$S\$K\$D\$M\$E\$V\$D\$V\$R\$L\$P\$F\$O\$A\$E\$M\$F\$O\$N\$Y\$I\$V\$F\$C\$G\$M\$A\$G\$Y\$P\$F\$E\$L\$V\$A\$V\$P	1020
D	b	961	K\$E\$F\$D\$T\$P\$T\$G\$R\$L\$N\$F\$S\$K\$D\$M\$E\$V\$D\$V\$R\$L\$P\$F\$O\$A\$E\$M\$F\$O\$N\$Y\$I\$V\$F\$C\$G\$M\$A\$G\$Y\$P\$F\$E\$L\$V\$A\$V\$P	1020
Q	y	1021	V\$L\$V\$E\$S\$V\$H\$I\$V\$S\$R\$V\$L\$E\$K\$R\$L\$D\$N\$T\$O\$S\$P\$E\$L\$S\$H\$T\$S\$S\$O\$G\$A\$T\$H\$A\$N\$K\$Q\$E\$F\$L\$H\$X\$O\$E\$L	1080
D	b	1021	V\$L\$V\$E\$S\$V\$H\$I\$V\$S\$R\$V\$L\$E\$K\$R\$L\$D\$N\$T\$O\$S\$P\$E\$L\$S\$	

```

Db 1141 GLFQFTVRLASLETARFTPSVERINHYITLTSLIEAPARIKKNAKPSDFWQEEVEVFENAE 1200
QY 1201 RYREMLPYLVKAVSTPIPKKEKIGIVGRGTSGKSSIGNALFRLVELSGGCIKIDGVRI 1260
Db 1201 RYREMLPYLVKAVSTPIPKKEKIGIVGRGTSGKSSIGNALFRLVELSGGCIKIDGVRI 1260
QY 1261 IGLADLRKSLSTIPEPVLFGSGTVASNDLPNQITTEDOIMALETHHKKEIAQLPLKLE 1320
Db 1261 IGLADLRKSLSTIPEPVLFGSGTVASNDLPNQITTEDOIMALETHHKKEIAQLPLKLE 1320
QY 1321 SEVMNGNFSGEQQLCLICIRALLRHCKIILDEPATAAMPTEDLLQETITREAFADCT 1380
Db 1321 SEVMNGNFSGEQQLCLICIRALLRHCKIILDEPATAAMPTEDLLQETITREAFADCT 1380
QY 1381 MLTIARHLTVAGSDRIWVLAQGQVVEEDTFESVLLSNDSSRFYAMFAAEKNKVAVK 1437
Db 1381 MLTIARHLTVAGSDRIWVLAQGQVVEEDTFESVLLSNDSSRFYAMFAAEKNKVAVK 1437

RESULT 2
US-09-001-273-2
: Sequence 2, Application US/09001273
: Patent No. 5994130
: GENERAL INFORMATION:
: APPLICANT: SHYUAN, Andrew
: TITLE OF INVENTION: NOVEL MULTIDRUG RESISTANCE-ASSOCIATED
: TITLE OF INVENTION: POLYPEPTIDE
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Testa, Hurwitz & Thibault
: STREET: 125 High St.
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/001,273
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: FENTON, Gillian M
: REGISTRATION NUMBER: 36,508
: REFERENCE/DOCKET NUMBER: ML-001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 248-7000
: TELEFAX: (617) 248-7100
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1453 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-001-273-2

Query Match 99.9%, Score 7298, DB 2, Length 1453;
Best Local Similarity 99.9%, Pred. No. 0;
Matches 1436; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKDDIGKEYITIPSGYSVRERTSTSGTHRDRESKFRFRPRPLECOALETFAAEGLS 60
Db 17 MKDDIGKEYITIPSGYSVRERTSTSGTHRDRESKFRFRPRPLECOALETFAAEGLS 76
QY 61 LDASHSOLRIIDEHPKGYHHGSLAKPIPTCKHHPVDNAGLFSCMTFSWLSLAR 120
Db 77 LDASHSOLRIIDEHPKGYHHGSLAKPIPTTCKHHPVDNAGLFSCMTFSWLSLAR 136
QY 121 VAHKGGEISMEDVMSLSKHSSSDVNCRRLERLMOBELNEVGDDASLRVYVIFCRRLI 180

```

```
|||||
Db 137 VAHKGELMEDVWSLSHRESSDVNCRLERLMQBELNEVGDAASLRVWVIFCRTLRI 196
Qy 181 LSIYVCLMTOLAGFSGPAPMKHLELYTOATESNLQYSLILVGLLLEIYVWSGLATW 240
Db 197 LSIYCLMTOLAGFSGPAPMKHLELYTOATESNLQYSLILVGLLLEIYVWSGLATW 256
Qy 241 ALNRTGRLGAILTMAFKLKLKNIKESLIGELINICSDGORMFEAAVGSLLAGG 300
Db 257 ALNRTGRLGAILTMAFKLKLKNIKESLIGELINICSDGORMFEAAVGSLLAGG 316
Qy 301 PVVALIGMIYVNIILGPTGLGSAVFIIFYPAMPASRLTAFFRRCKCAATDERQKNE 360
Db 317 PVVALIGMIYVNIILGPTGLGSAVFIIFYPAMPASRLTAFFRRCKCAATDERQKNE 376
Qy 361 VLTITKFKMTAWAKAFSOSVOKTREERRILEKAGYQOSTITGVAPIVVVIASVVTFSV 420
Db 377 VLTITKFKMTAWAKAFSOSVOKTREERRILEKAGYQOSTITGVAPIVVVIASVVTFSV 436
Qy 421 HMTIGFDLTAQAFTVTVVPMSPFALKVTPPSVKSLSSEASVAVDRFSLFIMEEVHMK 480
Db 437 HMTIGFDLTAQAFTVTVVPMSPFALKVTPPSVKSLSSEASVAVDRFSLFIMEEVHMK 496
Qy 481 NKPSAPRIKIEKNATLAMDSSHSISONSPLTPPKMKDKRASRGKKEKVRQLQRTHEQA 540
Db 497 NKPSAPRIKIEKNATLAMDSSHSISONSPLTPPKMKDKRASRGKKEKVRQLQRTHEQA 556
Qy 541 VLABQKHLILDSRPSPEEKEGHIHLGHLRLORTLHSDLELOEGKLVGICGSVSG 600
Db 557 VLABQKHLILDSRPSPEEKEGHIHLGHLRLORTLHSDLELOEGKLVGICGSVSG 616
Qy 601 KTSLSALIGOMTLEGSISATSGTAPYAAQOAMINATLRDNLLEKGEDEERYSVNS 660
Db 617 KTSLSALIGOMTLEGSISATSGTAPYAAQOAMINATLRDNLLEKGEDEERYSVNS 676
Qy 661 CCLRPDLALIPSSDLTEIGERGANISGGORORISLARALYSDRSIYIILDDPLSALDAHVG 720
Db 677 CCLRPDLALIPSSDLTEIGERGANISGGORORISLARALYSDRSIYIILDDPLSALDAHVG 736
Qy 721 NHIFSAIRKHLKSTVYLFVTHOLOLVDCDEVIPKESGCTERGTHERHMLNDIYATI 780
Db 737 NHIFSAIRKHLKSTVYLFVTHOLOLVDCDEVIPKESGCTERGTHERHMLNDIYATI 796
Qy 781 FNNLLIGETPPEINRSKETSODKPGKTSIKKEKAVKREEGOLVLEKGGOS 840
Db 797 FNNLLIGETPPEINRSKETSODKPGKTSIKKEKAVKREEGOLVLEKGGOS 856
Qy 841 VPMVYGYVYIOAAGPLAFLVYLMALFMLNVGSTAFSTWMLSYWIKOGSGNTVTRGNETS 900
Db 857 VPMVYGYVYIOAAGPLAFLVYLMALFMLNVGSTAFSTWMLSYWIKOGSGNTVTRGNETS 916
Qy 901 VDSKKNPDMHOYATIALSMAVWLILKAIRGVYFKGTLRASSRLHDELFRILRSPM 960
Db 917 VDSKKNPDMHOYATIALSMAVWLILKAIRGVYFKGTLRASSRLHDELFRILRSPM 976
Qy 961 KEFDTPTGRIILNRSKMDDEVYVRLPFOAEMFIONVILVFCVGMINGVFPWFLVAVGP 1020
Db 977 KEFDTPTGRIILNRSKMDDEVYVRLPFOAEMFIONVILVFCVGMINGVFPWFLVAVGP 1036
Qy 1021 LVILFVSLHIVSRVILRELKRLDNTIGSPFLSHITSSIOGLATTIHAUNKGOEFLRHYOEL 1080
Db 1037 LVILFVSLHIVSRVILRELKRLDNTIGSPFLSHITSSIOGLATTIHAUNKGOEFLRHYOEL 1096
Qy 1081 LDDNAPPEFLFCARMLAVRLDLSIALITTTGMLIVLMHQITPPAYAGLAISAVOLT 1140
Db 1097 LDDNAPPEFLFCARMLAVRLDLSIALITTTGMLIVLMHQITPPAYAGLAISAVOLT 1156
Qy 1141 GLFOFTVRLASTEARFTSVERINHYITLSLEAPARIKNKPSDPMOEGEVTEFENAE 1200
Db 1157 GLFOFTVRLASTEARFTSVERINHYITLSLEAPARIKNKPSDPMOEGEVTEFENAE 1216
Qy 1201 RYREMLPLVLKVSFTIKPEKIGIVGRGSGSSIGNALFRLVELSGGCIRIDGVRI 1260
Db 1260 RYREMLPLVLKVSFTIKPEKIGIVGRGSGSSIGNALFRLVELSGGCIRIDGVRI 1260
```

```
Db 1217 RYREMLPLVLKVSFTIKPEKIGIVGRGSGSSIGNALFRLVELSGGCIRIDGVRI 1276
Qy 1261 IGLADRSKLSITPEPVLFSSTVANSNDPFOYTEDQIMDLERTHKECIAOLPLKLE 1320
Db 1277 IGLADRSKLSITPEPVLFSSTVANSNDPFOYTEDQIMDLERTHKECIAOLPLKLE 1336
Qy 1321 SEVMNGDNFSVGEQOLICARALLRHCKIILDEATPAMDETDLLOETIREFADCT 1380
Db 1337 SEVMNGDNFSVGEQOLICARALLRHCKIILDEATPAMDETDLLOETIREFADCT 1396
Qy 1381 MLTIAHRLHTVLSDRIMVLAQGVVEFTPVSLLNDSRFPYAMFAAENKVAVK 1437
Db 1397 MLTIAHRLHTVLSDRIMVLAQGVVEFTPVSLLNDSRFPYAMFAAENKVAVK 1453

RESULT 3
US-08-843-459A-2
: Sequence 2, Application US/08843459A
: Patent No. 6162616
: GENERAL INFORMATION:
: APPLICANT: SHYUN, Andrew
: TITLE OF INVENTION: NOVEL MULTIDRUG RESISTANCE-ASSOCIATED
: TITLE OF INVENTION: POLYPEPTIDE
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESS: LAHIVE & COCKFIELD, LLP
: STREET: 28 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/843,459A
: FILING DATE: 16-Apr-1997
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Hanley, Elizabeth A.
: REGISTRATION NUMBER: 33,505
: REFERENCE/DOCKET NUMBER: MNI-056 (formerly MIL-001)
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)742-4214
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1453 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-843-459A-2

Query Match 99.9% Score 7298: DB 4; Length 1453;
Best Local Similarity 99.9%: Pred. No. 0;
Matches 1436; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKDIDIGKEYIIPSPGYSVRERTSTGTHRDREDSKFRRTPELCODALETAARAEGLS 60
Db 17 MKDIDIGKEYIIPSPGYSVRERTSTGTHRDREDSKFRRTPELCODALETAARAEGLS 76
Qy 61 LDASHMSOLRIIDEEHPKGYHGLSALKPIRTTCKHQHPVDNAGLFSQMTFSWLSLAR 120
Db 77 LDASHMSOLRIIDEEHPKGYHGLSALKPIRTTCKHQHPVDNAGLFSQMTFSWLSLAR 136
Qy 121 VAHKGELMEDVWSLSHRESSDVNCRLERLMQBELNEVGDAASLRVWVIFCRTLRI 180
Db 137 VAHKGELMEDVWSLSHRESSDVNCRLERLMQBELNEVGDAASLRVWVIFCRTLRI 196
Qy 181 LSIYCLMTOLAGFSGPAPMKHLELYTOATESNLQYSLILVGLLLEIYVWSGLATW 240
Db 240 LSIYCLMTOLAGFSGPAPMKHLELYTOATESNLQYSLILVGLLLEIYVWSGLATW 240
```

```
|||||
Db 197 LSTVCLMTOLAGSGGAFWVKHLETTQATNESNLQYSLLVLELLLETVRSNLSLWTW 256
QY 241 ALNRYTVRLGALITTAFFKILKLNKEKSLGELLINCSNDGQRFMAAASVSLAGG 300
Db 257 ALNRYTVRLGALITTAFFKILKLNKEKSLGELLINCSNDGQRFMAAASVSLAGG 316
QY 301 PVVAITLGMIVNIIIGPTGGLGSAVFLFFTPAMFASRLTAFFRRCKVAATDERVQKME 360
Db 317 PVVAITLGMIVNIIIGPTGGLGSAVFLFFTPAMFASRLTAFFRRCKVAATDERVQKME 376
QY 361 VLVYIKFKIKNYAWKAFSOSQKIREEBRILKEAGFQOSTVGAPlVVVIAVTFPSV 420
Db 377 VLVYIKFKIKNYAWKAFSOSQKIREEBRILKEAGFQOSTVGAPlVVVIAVTFPSV 436
QY 421 HMTLGFDTLAAQAFVTVTVNSMTFALKVTFEVSXSLSEASVADRFKSLTMEEVHMK 480
Db 437 HMTLGFDTLAAQAFVTVTVNSMTFALKVTFEVSXSLSEASVADRFKSLTMEEVHMK 496
QY 481 NKPASPIKITEMKNATLAMDSSHSSSIONSPKLPKMKDKKRASGKKKRYVROLQTEHOA 540
Db 497 NKPASPIKITEMKNATLAMDSSHSSSIONSPKLPKMKDKKRASGKKKRYVROLQTEHOA 556
QY 541 VLAQOKHLLDSDERSPEEBEGKHILGHLRLQRTLHSDLEIOEGKLVIGIGSVGSG 600
Db 557 VLAQOKHLLDSDERSPEEBEGKHILGHLRLQRTLHSDLEIOEGKLVIGIGSVGSG 616
QY 601 KTSLSIAILGOMTLLECSIASISGFAYVAAOAWILNATLRDNIIFGKREYDEERNVSLNS 660
Db 617 KTSLSIAILGOMTLLECSIASISGFAYVAAOAWILNATLRDNIIFGKREYDEERNVSLNS 676
QY 661 CCLRPDLAIISSDLEITGERGANLGGQRORIISALARYSDRSIYILDDPLSLDAHV 720
Db 677 CCLRPDLAIISSDLEITGERGANLGGQRORIISALARYSDRSIYILDDPLSLDAHV 736
QY 721 NHIFNSAIRKHKSKTYLFTVHOIQLYVDCDEVLFMEKGGCTTEGTHHEMLNLSGDVATI 780
Db 737 NHIFNSAIRKHKSKTYLFTVHOIQLYVDCDEVLFMEKGGCTTEGTHHEMLNLSGDVATI 796
QY 781 FNNLLGETPPEVEINSKKTSGSQKSDQKPKTSGIKKEKAVKPEEGQOLVQLEKSGS 840
Db 797 FNNLLGETPPEVEINSKKTSGSQKSDQKPKTSGIKKEKAVKPEEGQOLVQLEKSGS 856
QY 841 VPMSVYGVYIOAAGGPIAFVLYALFMLNNGSTAFSTWMLSYTKKSGSGNTTVRGNETS 900
Db 857 VPMSVYGVYIOAAGGPIAFVLYALFMLNNGSTAFSTWMLSYTKKSGSGNTTVRGNETS 916
QY 901 VSDSMKDNPHQYVASYIALSMAYMLILKAIKRGVYFVKGTIRASSRLHDELFRRLNSPM 960
Db 917 VSDSMKDNPHQYVASYIALSMAYMLILKAIKRGVYFVKGTIRASSRLHDELFRRLNSPM 976
QY 961 KFPDTPPTGRLINFRSKDMDEVDVRLPFOAEMFIQNVILVEFCVGMILAGVPMFLVAVGP 1020
Db 977 KFPDTPPTGRLINFRSKDMDEVDVRLPFOAEMFIQNVILVEFCVGMILAGVPMFLVAVGP 1036
QY 1021 LVILFSVLHIVSRKLEKRLDNTOSPFSLSHITSSIOGLATTHAYNKGOEFLHRYOEL 1080
Db 1037 LVILFSVLHIVSRKLEKRLDNTOSPFSLSHITSSIOGLATTHAYNKGOEFLHRYOEL 1096
QY 1081 LDDNOAEPFLPTCMRMLAVRLDILSTALITTTGIMLVIMHGOIPPVAVAGIAISYAVOLT 1140
Db 1097 LDDNOAEPFLPTCMRMLAVRLDILSTALITTTGIMLVIMHGOIPPVAVAGIAISYAVOLT 1156
QY 1141 GLFOFTVRLASETFARTSYERINHYIKTISLEAPARIKKNAPSPDPOEGEVTFFENAEM 1200
Db 1157 GLFOFTVRLASETFARTSYERINHYIKTISLEAPARIKKNAPSPDPOEGEVTFFENAEM 1216
QY 1201 RYRENLPVLVAKKVSFTTKPKKIGIVGTSGKSSLSGMALFRLVELSGCCIKIDGVARSID 1260
Db 1217 RYRENLPVLVAKKVSFTTKPKKIGIVGTSGKSSLSGMALFRLVELSGCCIKIDGVARSID 1276
QY 1261 IGLADLSKSLIIPQEVPLVSGTVRSNLDPPNOYTEDOINDALERTHMKECIAOLPIKLE 1320
|||||
```

```
Db 1277 IGLADLSKSLIIPQEVPLVSGTVRSNLDPPNOYTEDOINDALERTHMKECIAOLPIKLE 1336
QY 1321 SEVENDGNFVSGBEROLLICARALLRCKIILIDEATPAADTETDLLIOETIREAFADCT 1380
Db 1337 SEVENDGNFVSGBEROLLICARALLRCKIILIDEATPAADTETDLLIOETIREAFADCT 1396
QY 1381 MLTIAHRLHYLGSDRIMVLAQGVVEFDPSVLLSNDSSFTYAMFAAEKKAHVKG 1437
Db 1397 MLTIAHRLHYLGSDRIMVLAQGVVEFDPSVLLSNDSSFTYAMFAAEKKAHVKG 1453
```

```
RESULT 4
US-08-463-092B-4
; Sequence 4, Application US/08463092B
; Patent No. 5766880
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,092B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 08-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1546
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-6853
; TELEFAX: (613) 545-2342
; INFORMATION FOR SRO ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-463-092B-4

Query Match 28.5%; Score 2082; DB 1; Length 1531;
Best Local Similarity 34.5%; Pred. No. 1.8e-195;
Matches 488; Conservative 262; Mismatches 479; Indels 184; Gaps 21;
```


Qy	100	PVAGLESCMTSWLSLARVHKHKGELSMEDWVSLSHKSDVYNCRRRIHQEOLNE	155
Db	209	PESSARLSRTTWMITGLI-VAGYRQPLESGDLSLWLNKEDTSEQVYVPLVKMMKBECAK	267
Qy	160	V-----GPDA-----SLRWVJFCRTRIL	181
Db	268	TRKRPVAVVYSSKDPAPQKRESSVRDANEVEVLLYKSPQKKNPNSLFVLYKKTGPPTLM	327
Qy	182	SYVCLMTQOLAGFSGPAFWVKHLEEYTOATESNLQYSLLVJGLLLETIVRSMLATWA	241
Db	328	SEFFKAIHDMFMFSGPO-ILKILIKFVNDTKAPDMQGYEVYVLLFVTAOLQTLVHQFH	386
Qy	242	LNRYGTGRLAGALLTMAFKKILKLNKIRKS--LGELINICSDQGRPEAANGSLLAG	299
Db	387	ICFSGKRITAAVIGAVYRRKALVTFSARKSSTVGEIVNLMVSDAQRPMIDLATYINMTS	446
Qy	300	GPVVALIGMTYVNIIGPFGFSGNAFILIIFYBAMPASLTAYFRKKVATDEYQOKM	359
Db	447	APLOYILALVLLMLNTPPSVLAVAAVMLPVPNAVMAKTTIYVAAHMSKDNKRIKLMN	506
Qy	360	EVLVYIKFIMVAVMAVAFSOSYOKIEEERRIILERAGYFOSITTVAVPILVVIVVAS	419
Db	507	EILNGIVALKYAMELAFKOKVATINQDEELKYKKASAYLSANGTFTWCTPRIVALCTFA	566
Qy	420	VHMTLGD--LTAQAFTVYVYFNSKTFALKVTPSVKSLSASAIVADRFKSLFMEEVH	477
Db	567	VYVVIDENNITLDQTFVFSLATEINILRPLNLIPLWYISSIYASVSLKRLRFLFSHELE	626
Qy	478	--MKKNPASP---HIKIEKNATLAMDSSHSSIONSPLTPFKMKKDKRASGKKEKXRO	532
Db	627	POSTERRPVYDGGGTSITVKNATFTW-----	653
Qy	533	LQRTHOAVLEAQKHLLDSDRPERPEEGBEKHHLGLRLQTLHSDILETOEGKLVG	592
Db	654	-----ARSDP-----TLNGTTSIRPGALYA	675
Qy	593	ICGSVSGGKTSLSAILGQMTLLEGSIAISGTPRAYVAQOAMLTANDNLTGKEVDEE	652
Db	676	VVGVGCGGKSSLLSALIAEWDKVEGHVAIKGSAVYVPOQAMQNDLSRENILFGCOLEBP	735
Qy	653	RYNVLVNSCCLRPDLILPSDLTETGEGGANVSGGORISLARALVSDRSITYLDPL	712
Db	736	YKRSYVQACALLPDLLETILPSGDTLETGEGVNLSSGQOKORVSLARAVYSNADITYLDPPL	795
Qy	713	SALDAVGNHIFNSAI--RKHLKSTVLEFVTHOLOVLDQCDVIFPMKBCGCTERTGTHEEL	770
Db	796	SAVDAHGKHIFENVIGPKMKLNKTRILTYHSMSTLPQVDVITYMSGKSIEMKSYOEL	855
Qy	771	MNLNGDYATTF-----NNLL-----GETPVEIN	795
Db	856	LARGGAFAEFLRTYASTEDQDAEENGVTYGVSGPCKEAKOMENGMLVYDASGLOQROLS	915
Qy	796	SKRTTSOSQKKSDQKGTGKYSIKKKEKAVPREGOLVQLEEKQSVSPSYVGYIOAGG	855
Db	916	SSSSYSDDISRHHN--STAELOKAAKKEETWKLMEAKQOTGYVKLSYWDYKRAIGL	972
Qy	856	PLAFLVLMALFMLNVGSTASTWMLSYWKIKQSGAGTTVYRGNETSVSDMKNPIMOYYA	915
Db	973	FISTLSIF-LPMCNHVSALASNWYLSMTWDDPIVNT--QEHTRYVRSYVGCALGISOCI	1028
Qy	916	SYVALSAVMMLIKAIRGVYVKGTLAASSRLHDELFRRIILRSPKPFDTPTGTILNRF	975
Db	1029	AVEFYSAVNSI-----GGILASRCLHYVLDLSILRSPSPFERPESGNLVNRF	1076
Qy	976	SKDDEVDVRLPRQAEKFIONVLYFVCYGMLAGVPMPLVAVGGLVILFSLVLIHVSVL	1035
Db	1077	SKELDYDSMTPYKIKFMKESLENVIGACTIVLATPRIAIITLIPGLIY--FVQGFY	1133
Qy	1036	I---RELKRDNIOTSPFSLHITSIOGLATIHAYKKQGEFLHARQELLDNQAFPLFT	1092
Db	1134	VASSRQJLRKLESRSRVSYHFMETILGVSVIRAFEEQGRFLHQDGLKVDENQKRYATSI	1193
Qy	1093	CAMRWLAVRDLISIALITTTGMLIVLHNGQIPRAYAGLAISYAVOLTGLEQFQTVRIASE	1152

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	28.5%	Score 2082	DB 2	Length 1531
Best Local Similarity	34.5%	Pred. No. 1.8e195		
Matches 488	Conservative 262	Mismatch 479	Indels 184	Gaps 21
QY	100	PVDNGLPSCMTFMWSSLSLARVAHKKGELSMEDVSLSKSHESDVNCRLERTLMOEELNE	159	
Db	209	PSSASFSLRITFPMWITGLI--VGRGRPLGSGDMSLNKEDISBDQVPLVKNMKKECK	267	
QY	160	V-----GPDA-----SLRVRVIFCRTLIL	181	
Db	268	TRKQPKVYVSSKDAQAPKRESSKYDANEVEALIVKSPQENPSLFKLYTFTGPEYLM	327	
QY	182	STVCLMTIQLAGSGPAPFWKHLLEYTOATESNLQYSLLLVGLITELTVRSMSLATFWA	241	
Db	328	SFEFAHIDLMFSGPO--ILKLLIFVNDTKAPDMOGFYTYLLVYTAQLQTLVLYHQYH	386	
QY	242	LNRYTVBLRGAILTMARFKTILIKNIKES--IGELINICSDGQRMFEAAVGSLLAG	299	
Db	387	ICFVAGMKRIKTEVAVYAKVRLVITNSARKSSSTGEIVELMSVDQRFMDLATVYINMIS	446	
QY	300	GPVVALIGMIYVILITLPGFGSAVFILFYFAMFASLTLYFPFRKCAALDERQCKN	359	
Db	447	APLOYILALYLLMLVGLSVLAGVAVMLVPMVNAVMARKITYAVAHKSKDNRIKLMN	506	
QY	360	EVLATIKFIKMYAKFASQSYOKIREERIRILEKAGFYQSITVGVAPIVVIASVYFS	419	
Db	507	ELNIGIKYLLKYLAWELARKDYLAIROEELKALKKSATILSANGFTWCTPPLVALCFPA	566	
QY	420	VHMTLGFED--LTAQAFVTVVYFNSKTRALKYTPRSVKSLSASAVANDRFKFLIMEYH	477	
Db	567	VYVITDENNIIDAQAFPSLALFNILREPLNPLMVISSIVQASVSLKRLRFLFHEELE	626	
QY	478	--MIKRPASP---HIKTEKNATLAMPSSHISQNSPLPLPMMKKDKRASGKKKEKYNQ	532	
Db	627	PDSIRRRVYKDGGSNSTITVRNATITW-----	653	
QY	533	LQRTHQAVLAEQKHLLDSDERSPEEELKNIHLGLRLQRTLHSIDLEIOEGKIWG	592	
Db	654	-----ARSDPP-----TLNGITFSIPGALVA	675	
QY	553	ICGSVGSKRTSLISAILDOMTLBESIASGFAFVAQAQAILNATLRNILFGKEYDE	652	
Db	676	VVGQGGCKSKSLSLALLEMKVBECHVAIKGSVAAPQAWIQONSLRILNITLFGQLEEP	735	
QY	653	RYSVYVNSCCLEPDLAIIPPSSDLTEIGERGAMLSGSGORSLSLARALYSRISYILDPUL	712	
Db	736	YRSVYIQACALLPDELIIPLSGDRTIEIGERGAVLSSGQQRVSLARAAYVSNADYILEDPUL	795	
QY	713	SALDAHVGNIHFNSAI--RKHLKSTVLEFYTHOLOLYLDCDEVITFMKESCTIERGTHEEL	770	
Db	796	SAVDAAHVCKHIFENYIGFRGKMLKNTRIILVTHSMYSYLPQVDVITVMSGKISEMSYOEL	855	
QY	771	NMLNDVYATIF-----NNLL-----GEPFVEPVEN	795	
Db	856	LARDQAFLEPLRTYASTQEDQDAENGVTVYSGPKEKAKOMENGMALVYDSAGKQLOARQIS	915	
QY	796	SKKETSGSQKSSQDKGPKTGSIRKKEKAVKPEBGOLVOLEEKQSGSVPSVYGVYIOAGG	855	
Db	916	SSSSVSGDISRHHN---STALQKAEAKKEETWIKLMEADKAQTGVQKLSVYDYDKAICGL	972	
QY	856	PLATVYALMLFMNLVNGSTAFSTWMLSYWIKQSGGNTTYTRGNETSVSDSKMKNPMQIYA	915	
Db	973	FISFSLIF--LEMCNIVSALASNYWLSLTWDTDEIVNGT---QERTVYRLSVYGALISOGI	1028	
QY	916	SIYALSNVYMLILKARGVVEFKGTALRSSRHLDELFRILSSPKMFPDPTPTGRIINF	975	
Db	1029	AVFEGSMAVSI-----GGILASRCHLDVLLHSILRSPSIFERTSGNMLVNRF	1076	
QY	976	SKMDQEDVDRLEFQAEFIQYNIIVLFCVGMATGVFPWEFLVAVGRLVILFSLHIVSRVL	1035	
Db	1077	SKEDTVDSMKPEVYIKMKGSLFNVIAGACIVILATPTAAILIIPGLIGLY---FFVQGRY	1133	

QY 1036 I REKRLDNTOSBPFSHTISSIOGATTHAANKGOEFLHRYOELLIDNOAPFLELT 1092

Dd 1134 VASSRQLRKLESRSRPVSHENETLLSVSLRAFEBOERFHOSDLKVDEOKAYTPTI 1193

QY 1093 CAMWLAVLRDILSIALITTTTGIMIVLMAGQIPPAVYAGLTAISYAVALUTGLFOFTVRASE 1152

Dd 1194 VANMWLVAIRECGNOCVILEAFALFAVISHSLSAGLVGSYSYLQVTYTTLINMLYBMSSE 1253

QY 1153 TEARFISVERINHKTLSLEAPARKTKRAPSPDMPOBGEVTFENAEMKRIRBNLEPLVUK 1212

Dd 1254 METWIVAVEKLEKESSET-EKEAWQIOELFAPPSSWPQVGREVEFRNYCRRYREDDFPVLRH 1312

QY 1213 VSPIFKERKIGVGRGSKSSLGMAFLRVLESGCCIKIDGVHISDIGADLSKLISI 1272

Dd 1313 INVTTNGCEVGVIGRTGACKSSLTGLTRINESABGELLIIDGINAIKIGLDLRFKTTI 1372

QY 1273 IPOEPVFEGSTVSNDLPFNQYEDOIDMALERTHMKECIAOLPLKLESEMENGDNFSV 1332

Dd 1373 IPQOPVLFSSLSMNNDPFQSYDEVMWSLELAHKDPVSALPKRLDHCEAGESENLSV 1432

QY 1333 GEROLLICIAALLHCKKILILDBATAAMDTEFDLLIQETIRAFADCTMLTIABRLFTVL 1392

Dd 1433 GORLOVCIALARALKRKRIIVLDEATAVPLEFDLLOSTIRFOFEDCYVLTIAHLRNLTIM 1492

QY 1393 GSDRIMVLAAOGVPEPTPSYSDNISNSRFFYM 1425

Dd 1493 DYTRVIVLDKGELQETGAPSDDL-QQRGLFTSM 1524

RESULT 6
US-08-460-907B-4
Sequence 4, Application US/08460907B
Patent No. 5891724
GENERAL INFORMATION:
APPLICANT: Deeley, Roger G.
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
NUMBER OR SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PARTO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
City: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:

NAME: Steeg, Carol Miernicki
 REGISTRATION NUMBER: 39,539
 REFERENCE/DOCKET NUMBER: 01551
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (613) 545-2342
 TELEFAX: (613) 545-6853
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1531 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-460-907B-4

Query Match 28.5%; Score 2082; DB 2; Length 1531;
 Best Local Similarity 34.5%; Pred. No. 1,8e-195;
 Matches 488; Conservative 262; Mismatches 479; Indels 184; Gaps 21;

100 PVDNAGLSCMTFWLSLAVAHKKGELSMEDWVLSKHSSDVNCRRLERLMOEELNE 159
 209 PESSASFLSLTFEWMITGLT-VRGTRQPLEGSDMLSKEDTSEGVVPLVKNMKKECAK 267
 160 V-----GPPAA-----SLRRVWIFCTRLIL 181
 268 TRKQPVKVVYSSKDPQKRESSKYDANEVEALIVKSPQKEMNPFLFVLYKTFEPFLM 327
 182 SIVCLMTQLAGFSGPAPVAKHLEETQATESNLOSLVLLGLLTETIVRSWLSALTWA 241
 328 SFFFAKIDLMFAGPQ-ILKLLIKFVNDTKAPDQGYFYVLLTFTVACLOTLVLYHQYFH 386
 242 LNYRQVRLGAILLMARKLILKLNIEKS--IGELINICSDQORFEEAAGVSSLAG 299
 387 ICFVGMKIKTAIVGAVRKALVLTNSARKSTVEIYNLMSVDQRPMDLATTYNNMWS 446
 300 GPVAILGMIYVNIITGPTGELGSAVFLFYPPAMFASRLTAFFRKCAVATDERVOKMN 359
 447 APLOYTIALYLLMLNGLPSVLAVAVWLVAVPVNAVMAMKFTTYVAHMKSKDNKIKLMN 506
 360 EVLYIKIKIKYAVYKASOSVOKIREBERILKEAGYFQSTITGVAVPIVVIVIASVTFPS 419
 507 EILNGIKYKLYAMELAKRDKVLAIRQELKYLKKSAYLSAAGTPTWCTPPLVALCTFA 566
 420 VHMILGFD--LTAQAFVYVVFVNSMTFALKYTPSVASLSASAVYDREKSLFMEVH 477
 567 VYVITDENNIIDAGTAFSLALFLRPLMLLPVIVISSIVASVSLRLRFLSHHEELE 626
 478 --MIKNRPASP---HIKTEMKATLADSSSHSSSIONSPKLPFKMKDKRASRGKKEVRO 532
 627 PDSIFRRVYKDGCGNSTITVRNATFTW----- 653
 533 LQRTHEQAVLAQKGLHLLDSDERSPEEEDKHTHLGLRLQRTLSIDLEIQEGKLVG 592
 654 -----ARSDP-----TLNGITFSIPGALVA 675
 593 ICGVSGKTSLSAILQOMTLBESIAISGTFAYVAQOATLANTLRDNLIFGEYDEE 652
 676 VVGQVCGCKSSLLSALAEKMKVEGHVAIKGSVAVPOAMQONSLRENTLIFGCOLERP 735
 653 RYNSVJNSCCRLPDLAIPSSDLTEIGERGANLGGORORISLARATSDRSIGIYLDPL 712
 736 YRSYVQACALLPDLLEIPSGDRTEIGERGVNLSSGOKORSALAAVSNADIIYLFDDPL 795
 713 SALDAHVGNHIFNSAI--RKHLKSTVLEFVTHQOLYVDCDEVIFMKBGCTIERGTHEEL 770
 796 SAVDAHVKGKHFENYVIGKMLKNTKRLIVTHSMGYLPQVDIIVMSGKISEMSGYEL 855
 771 MNLNGDVATIF-----NNLL-----GETPPEVN 795
 856 LARDGAFLEPLTYASTEOEDAENGVTGVSGPKKAKOMENGMLVYDSAKOLOROLS 915
 796 SKKETSQOKSODKPKTGSIKKEKAVKPEEGOLVQLEKQGSVSPMSVYVYIOAAGG 855

916 SSSYSXGDISRRHN---STAELQKAERKKEETWKLMEADKAQOTGVKLSVYDYNKAIGL 972
 856 PLAFVIALALEPLANGSTAFSTWMLSYWIKOGSGNTYTRNGETSVSDMKONPHMQYYA 915
 973 FTSFLSIF-LFMCNHNVSALASVWLSLMTDPIYNGT---QEHTVRLSVYGALCISQCI 1028
 916 SIYALSMVMTLKAIRGVVFEVFKTLRASSRLHDELFRRLRSLSPKFFDTPPTGRLNRF 975
 1029 AVFGSMAVSI-----GGILASRCLHVDLHLSILRSPMSFPERPSSMLVNRF 1076
 976 SKMDDEVVRLPEQAEFTQNIYVLFPCVGMIAVGPMPVLAAGVPLVILFSLHIVSRYL 1035
 1077 SKELDTVDSMIEPVYIKMKMGSIFLVNIGACIYVLLATPIAIIIPPLGIIY---FVQREY 1133
 1036 I---RELKRDNIQSPFLSHITSSIOGLATHAANKOEFLHRQELLDDNOAFELFT 1092
 1134 VASSQOLRLESVSHSPVYSHNETLLGAVYIRAFEEQERFHOQSDLVADENQKAYPSI 1193
 1093 CAMRLAVRLDISALITTTGLMIVLMHGOIPPAVAGIATSAVQULGTFQFVRLASE 1152
 1194 VANRLAVRLKEVCNCIYLFALFAVIRSHLSAGLVGLSVSYQVTTYLWMLVYRMSSE 1253
 1153 TEARTSVYRINHYIKTSLLEAPARIKKNKAPSPDPQEGEVTFFENAEKRYRENLPVLYRK 1212
 1254 METNIVAEVRLKEYSET-EKEAPMOIQETAPPSWPQYGRVFEFRYKCLYREDLDFVLH 1312
 1213 VSTFKPEKIGIVRTSGKSLGMALEFRLVELSGGCKIDGVNISDGLADLSKLSI 1272
 1313 INVTINGEKVGIVRTGAKSLTFLGRINESAGEETIIDGINAKIGLHDLFFKTI 1372
 1273 IPOEVPLFSGVYRSLNLPDNOFTEOIMDALERTHMKECIAOLPLKLESEVWENGDNFSV 1332
 1373 IPDVPYFSGSRIMNLDPPSOYSDSEVWTSLELAHLKDVSALPKLDHECEGGENLSV 1432
 1333 GEROLLACIALRLRCKLILIDEATAAMDTEDDLIOETIREAFADCTMLTAHRLATVL 1392
 1433 GORQLVCLARALLRKTILILIDEATAAVDLETDLIQSTIRFQEDCFVLTAFHRLNTIM 1492
 1393 GSDRTMVLAAQGVFEFDPISVLLSNDSSRFYAM 1425
 1493 DYTRVIVLDKGEIOEYGAPSDLL-QOORGLFYSM 1524

RESULT 7
 US-08-463-179A-4
 ; Sequence 4, Application US/08463179A
 ; Patent No. 6001563
 ; GENERAL INFORMATION:
 ; APPLICANT: Cole, Susan P. C.
 ; APPLICANT: Dealey, Roger G.
 ; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street, suite 510
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/463.179A
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/966,923
 ; FILING DATE: 27-OCT-1992
 ; APPLICATION NUMBER: 08/029,340
 ; FILING DATE: 8-MAR-1993

```

? APPLICATION NUMBER: 08/141,893
? FILING DATE: 26-OCT-1993
? APPLICATION NUMBER: 08/400,207
? FILING DATE: 20-MAR-1995
? ATTORNEY/AGENT INFORMATION:
?
? NAME: DeConti, Giulio A. Jr.
? REGISTRATION NUMBER: 31,503
? REFERENCE/DOCKET NUMBER: POI-002CP8
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 227-7400
? TELEFAX: (617) 227-5941
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1531 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? OS-08-463-179a-4

```

Query Match	28.5%;	Score 2082;	DB 3;	Length 1531;
Best Local Similarity	-34.5%;	Pred. No. 1.8e-195;		
Matches 488;	Conservative 262;	Mismatches 479;	Indels 184;	Gaps 21;

QY	100	PVDNAGFSCSTFEMLSLRAVAKKELSDMEDVYSSKHESSOVNCRLEIRMOELNE	159
Db	209	PSSASFLSRTTFMWITGLI-VKROYRPLGSDLSLKNEDTSEQVVPVLVYKNNKCCAK	267
QY	160	V-----GPDAA-----	180
Db	268	TRKOPKVVYSSKPDPAQPKESSKVDANDEEVALIVKSPQKEMNPSTLKVYKTFGEFYFLM	327
QY	182	SIYCLMTTOLAGEGPRAPMKHLELYQATIESNIQYSLLYLGLLTIELYRSMSLATMA	244
Db	328	SEFFKALHDLMMFEFGPO-IKLKLKIFENDKRAPDMOGYFTYVLLFTRACLOTLYLHOYFH	386
QY	242	LNTRYGRALRCALITMAFKPILKIKLNKIKERS--LGEILINICSDNGOMPEAAVAGSLIAG	289
Db	387	ICFVSGHRKIRKAVGAYYRRALVYTTNSARKSSITYGELVNLMSVDQAQFMFLATYIMNIMS	446
QY	300	GPVVAIIIGMIYVNIILGPFGLSAVEFILEYPAMMFASRLTAYFRRCVATDERVOKMN	359
Db	447	APLQVITIALYLLMLNLPBSVLAVAVVWLAMPVAVAMAKTKTYQVAHMKSKDIRIKYLM	506
QY	360	EVLTYIFIKRYANVAKFASQSVQKIRREERILKAGYPOSITVGAPIYVVLASVTPES	419
Db	507	EILNGIKVLKLYAELEFKRKVLAIRDEELKVLKRSAYLAVGFTWCPTPVALCTFA	566
QY	420	VHMTLGFDP--LTAQAQFTVVYVNFMSMFALKVPFYSKSELSAADVADREKSLFEMEYH	477
Db	567	VYVVIDENNILLDAQTAVALSALFPIILFPLINLPMTVLSSTVQASVSLKRLRIFLSHELE	626
QY	478	--MIKNPASP--HIKIEKNATLAWDSHSSIONSPLTPPKMKDKRASRGKKEVRQ	532
Db	627	PSDIERPRVKDGGGTSIYVNAATFTM--	653
QY	533	LQRTERHOAVLAEQKGHLLSDDERPSEEBEGKHILGHLRLQRTLSIDLEIOEGLVG	592
Db	654	-----ARSDPP-----	679
QY	593	ICGSVSGSKTLLISAILIGOMTLEGSTAISTGTPYVAQOAMILNATLRDNLGCKEVED	652
Db	676	VWGGVGCGKSSLSLALLAEMDKVGVHAIKGSVAVYPOQAMTQDSDRENTILFECQLEEP	735
QY	653	RYNLSVNSCCRLPDLALIPSSDLTEIEERGANSLSGGORRISLARALYSRSITYLDDPL	712
Db	736	YKRSVIOACALIPDLLELIPSDRRIETEKGNLSGGOKQYVRSIARAYNSNADITYLPDPL	795
QY	713	SALDAVGNHIFNSAI--RKHLSKTYLVYTHQLOYLVDCDEVILFMKEGCTITERGTHIEL	770
Db	796	SAYDAVAGKHIEFENVIIGPKMLKTKRTILTYVHSMYSYLPQVDVITVMSGKRISEMGYOEL	855
QY	771	MLNNGDYATIF-----NNLL-----GETPPVEIN	795

[illegible]

RESULT 8
US-08-461-384B-4
Sequence 4, Application US/08461384B
Patent No. 6025473
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,384B
FILING DATE: 05-JUN-95
PRIOR APPLICATION DATA:

```

1      APPLICATION NUMBER: 07/966,923
2
3      FILING DATE: 27-OCT-1992
4
5      APPLICATION NUMBER: 08/029,340
6
7      FILING DATE: 8-MAR-1993
8
9      APPLICATION NUMBER: 08/141,893
10
11     FILING DATE: 26-OCT-1993
12
13     APPLICATION NUMBER: 08/407,207
14
15     FILING DATE: 20-MAR-1995
16
17     ATTORNEY/AGENT INFORMATION:
18
19     NAME: Steeg, Carol Miernicki
20
21     REGISTRATION NUMBER: 39,539
22
23     REFERENCE/DOCKET NUMBER: 01547
24
25     TELECOMMUNICATION INFORMATION:
26
27     TELEPHONE: (613) 545-2342
28
29     TELEFAX: (613) 545-6853
30
31     INFORMATION FOR SEQ ID NO: 4:
32
33     SEQUENCE CHARACTERISTICS:
34
35     LENGTH: 1531 amino acids
36
37     TYPE: amino acid
38
39     TOPOLOGY: linear
40
41     MOLECULE TYPE: protein
42
43     IS-08-461-384B-4

```

Query Match	28.5%;	Score 2082;	DB 3;	Length 1531;
Best Local Similarity	34.5%;	Pred. No. 1.8e-195;		
Matches 488;	Conservative 262;	Mismatches 479;	Indels 184;	Gaps 21

QY	100	PVDNAGLEFSCMTEFWLSSLARVAHKKELSMEDVWLSKHESSDYNCRLERLMOEPLNE	159
Db	209	PESSASFEJRTTEFWITGGL-VRGYROPLEGSDLMSLNKEDTBSQVVPVLYVKKMKCECA	267
QY	160	V-----GPAA-----SLRWVWICFRRLIL	181
Db	268	TRKQPVKVVYSSKQAPQKRESSKVDANEEBALVKSPOKEMNPJLEKVLKTFEGPEFLM	327
QY	182	SIVCLMTITOLAGEGPRAPMKHLEIYQATRESLQYSLLYGLLLEIYRSMSTALATNA	241
Db	328	SFFKRALHIDLMFMSGPO-IKLILKIFVNDKRAPDMOGYFYTVALLEFYACLOTVLHQFYH	386
QY	242	LNRYFVGLRGCAIITMAFPKILIKLNKIKERS--LGEILINICSDNGOFMEFAAVGSILAG	299
Db	387	ICFVSGMKIKAVIGAVYRKRALVITNSARKSSYTGVELVNLMSVDAQOFMDLAIYINIMWS	446
QY	300	GPVVALIGMIYVITILGPTGLGSAVITLTPPAMMFSLTAIFRRKCVAAATDERQOKAN	359
Db	447	APLOYIILALYLLMLNLPSPVLAGAVALVWLVPVAAVMAKMTKTYQVAAHMSKDNIRIKLM	506
QY	360	EVLVYIKPIKMYAWWAKFOSVOKIREEBRILIEKAGYPOSITVGAPIVVIVLASVYTPS	419
Db	507	EILNGIKYKLIYAMELAFKDKVLIAROBELVYLKKSATLSAGFTTWCPIPELVALDTFA	566
QY	420	VHMTLGED--LTAQAFTVYVTFNSMTFALKVTPFYSVKSLEASVAADRKSLFLMEVH	477
Db	567	YVYTIIDENNIIIDAOTAEFSLFLNITLFPILNIPMIVISSIYQASVSLKRLRIFLSHEELE	626
QY	478	--MIKNKRASP-----HIKEMKNATLAWDSSHSIIONSPLKLPKKKKRKRKRCKKEVRO	532
Db	627	PDSTERRRKPVDGGTNSITVYNAFTW-----	653
QY	533	LQRTHEQAVLAEOKGHLLDSDEPRPSEEBEGKHIHLGHLQRTLSHIDLEIOEGKLV	592
Db	654	-----ANSDDP-----TLNGITFISPEGLVA	675
QY	593	ICGSYGSKRTSLISAILGOMTLLEGSTAISCTFAYVAAQAMILNATLRDNLFEKEYDER	652
Db	676	VVGQVGCKRSLSLALAEKMDYEGHVAIKSVAVYVQAQAMIONDSRENLIFEGCOLDEP	735
QY	653	RNSYVJNSCCRLPDALIPSSDLIEIGRGNANSGGGRQORISLARALYSRSTYIILDDPL	712
Db	736	YRSTYIQACALLPDELILPSSDKREIEIGKYNLSGGQKQRYSLARAYYSNADITLFPDPL	795
QY	713	SALDAHVHNHINSAI--RKHLKSKTALVTVYHOLQYLVDCDEVIFEMKEGCTITERGTHEEL	770

```

Db 796 SAVDAHVGKILIFENVIGCPKGLKNKTRILVTHSMSTLPOVDYIITVMSGGKISEMGSGOEL 855
QY 771 MNLNGDVAITF-----NNLL-----GEPPEYEN 795
Db 856 LARDGAAVEAFRLIYASTDEQDAENGVTGVSQPGKEAKOMENGLVYDSAGKOLQORLS 915
QY 796 SKKETSSQKKSDQKGPKTGSIKKEAAVPEEGOLVQLEPEKQSGVPMYSYGVYITQAAG 855
Db 916 SSSSYSGDIRHNH--STAELQAKAKKEEYKLMLEADKAOTGVKLSYWDYWKAGLG 972
QY 916 SIYASMAVMLIKAIRGVYVKGTLRASSRLHDELFRILRSPMKFPDTPGTGILNRF 975
Db 1029 AVFGYSMAVSI-----GGILASRCLHVDLHLSILRSPMSFEPFRPGNLVRF 1076
QY 976 SKMDDEVDRILPROAEKFIQNTVLVFFCYGMLAGVPEWPLVAVGGLVILFSLVHIVSVL 1035
Db 1077 SKELDTVDSMIPVKKFMKSLNVLNAGACIVILATPIAAIIIPGLGITY--FFVOPEY 1133
QY 1036 I---RELKRLDNITOSPFLSHITSSIOGLATIIAHYKKGQEFLLHRYQELLDNOAPFFLT 1092
Db 1134 VASSRQKRLRESYRSRPSVYSHENETILGSYVIRAFPEQGRFHOQDLKVDENQKAYYSI 1193
QY 1093 CAMBLAVRLDLISALITTTGMIIVLMHGQIIPPAYAGIAISAVOLTGLFOQTVRLASE 1152
Db 1194 VANRLAVRLECGVNCIVLPAALFVVISRSLSAGLVGSYSGLQVTVYINLVLRBMSE 1253
QY 1153 TEARFTSYERINHYIKTSLSEAPARKKAPSPDMQOESEVTFEVAENKRYRENPLVLYK 1212
Db 1254 METNIVAVERLKEVSET-EKEAPWQIOETAPSPSWQVGRVFRNYCCRYRREDDDEVLRH 1312
QY 1213 VSFITKPEKRGIGYRFGSGKSSLGMAFLPVLETSGGCIKIDGVRISDGLADLSKLSI 1272
Db 1313 INVITNGEKVGIYGRGAKRSSTLTGLFRINESAAGEIIIDGINARIKIGLHDLRFKTI 1372
QY 1273 IPOBVLFSGTVSRNLDLPFQSGYDEEYVMSLELAHLKDEVSAALPKDLHECEGENSTV 1432
Db 1373 GEROLLCIARALLRHCKILLIDERTAAQMTENDLLQETIRRAFADCMTLTAHSLHYL 1392
QY 1393 GQORLVCLARLLKTKTILLVIDERTAAVADLETDDLQSTIRQFEDCTVTLTAHRLNTIM 1492
Db 1493 GSDRIWLAQOVVEFPTPSVLLSNDSSRPFYAM 1425
QY 1425 GSDRIWLAQOVVEFPTPSVLLSNDSSRPFYAM 1425
Db 1493 DYTIRIVLDKEIOETGAPSDLL-QQRGFLFYSM 1524

RESULT 9
US-08-141-893-2
Sequence 2, Application US/08141893
Patent No. 5489519
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, SUITE 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```


QY 1333 GEROLCTIARALLRHCKILILDENTAAMTETDILIOETIREAFADCTMLTIAHRLFTVL 1392
1333 GEROLCTIARALLRHCKILILDENTAAMTETDILIOETIREAFADCTMLTIAHRLFTVL 1392
Db 1433 GOROLVCLARALLRKTKILVLENTAAMTETDILIOETIREAFADCTMLTIAHRLFTVL 1492
1433 GOROLVCLARALLRKTKILVLENTAAMTETDILIOETIREAFADCTMLTIAHRLFTVL 1492
QY 1393 GSDRIWLAOGVAFETPSPVSLNDSREYAM 1425
1393 GSDRIWLAOGVAFETPSPVSLNDSREYAM 1425
Db 1493 DYTRVIVLDKEIOETIETGAPSDLL-QORGLFTYSW 1524
1493 DYTRVIVLDKEIOETIETGAPSDLL-QORGLFTYSW 1524
RESULT 13
US-08-463-179A-2
Sequence 2, Application US/08463179A
Patent No. 6001563
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deoley, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
City: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,179A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Decont, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQI-002CP8
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-179A-2

Query Match 28.4%; Score 2075; DB 3; Length 1531;
Best Local Similarity 34.5%; Pred. No. 9e-195;
Matches 487; Conservative 261; Mismatches 481; Indels 184; Gaps 21;

QY 100 PYDAGLFCSCMTESLRLARVAHKKGELSMEDVWSLSKSHSSDPVNCRRRLRLOEELNE 159
100 PYDAGLFCSCMTESLRLARVAHKKGELSMEDVWSLSKSHSSDPVNCRRRLRLOEELNE 159
Db 209 PESSASLSTRTTEFWITGLI-VRGYROPRLSGSDLSLNKEDTSQVVPVLVKNKKKCEAK 267
209 PESSASLSTRTTEFWITGLI-VRGYROPRLSGSDLSLNKEDTSQVVPVLVKNKKKCEAK 267
QY 160 V-----GPDAA-----SLRRVWIFCRTLIL 181
160 V-----GPDAA-----SLRRVWIFCRTLIL 181
Db 268 TRKOPKVVVYSSKRPAPKRESSKVDANEVEALIVKSPQKEMNSLKKVLYKTFGPFLM 327
268 TRKOPKVVVYSSKRPAPKRESSKVDANEVEALIVKSPQKEMNSLKKVLYKTFGPFLM 327
QY 182 SYVCLMTITQLAGSFGAPFPMVKHLETTQATESNLQYISLLVGLLLEIYVNSGLALTWA 241
182 SYVCLMTITQLAGSFGAPFPMVKHLETTQATESNLQYISLLVGLLLEIYVNSGLALTWA 241

Db 328 SFFPKAHLDMFSGPD-LIKLKLFVNDTKAPDMQCYFTVLLVFTYACLOTIVLHYEFH 386
328 SFFPKAHLDMFSGPD-LIKLKLFVNDTKAPDMQCYFTVLLVFTYACLOTIVLHYEFH 386
QY 242 LMYETVRLRGALITVMAFKKILKLNKIKES--LGELINICSDGORMFEAAVGSLLAG 299
242 LMYETVRLRGALITVMAFKKILKLNKIKES--LGELINICSDGORMFEAAVGSLLAG 299
Db 387 ICFVSGMRITAVIGAVYRALKALVTNSARKSSVYGEIVNLMSVDAQRFMDLATYINNIMS 446
387 ICFVSGMRITAVIGAVYRALKALVTNSARKSSVYGEIVNLMSVDAQRFMDLATYINNIMS 446
QY 300 GPVVALLGMYVNYLLCPPTGLSAVIFLFPYPMMEFSLRTAFRRKCVANAATERVOKAN 359
300 GPVVALLGMYVNYLLCPPTGLSAVIFLFPYPMMEFSLRTAFRRKCVANAATERVOKAN 359
Db 447 APIQVILALYLLMLNLPVLAGVAVVLMVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 506
447 APIQVILALYLLMLNLPVLAGVAVVLMVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 506
QY 360 EVLTYIKETMYAVKAFSOSVOKIREERRILEKAGYFOSITGVAPVIVVAVVAVVFS 419
360 EVLTYIKETMYAVKAFSOSVOKIREERRILEKAGYFOSITGVAPVIVVAVVAVVFS 419
Db 507 EILNGIKVLYKLYAMELAFKQVLAIRQELKVLKSKAYLSAVGFTWCPPLVALCTFA 566
507 EILNGIKVLYKLYAMELAFKQVLAIRQELKVLKSKAYLSAVGFTWCPPLVALCTFA 566
QY 420 VHMTHGPD--LTAQAFVTVVNSMTFALKVTPFSYKSLSEASVAVDRKSLFLEEVH 477
420 VHMTHGPD--LTAQAFVTVVNSMTFALKVTPFSYKSLSEASVAVDRKSLFLEEVH 477
Db 567 VYVVIDENMLDQATAVSLAFNLIRPLNLPVIVISIVQASVSLKRLRIFLSHELE 626
567 VYVVIDENMLDQATAVSLAFNLIRPLNLPVIVISIVQASVSLKRLRIFLSHELE 626
QY 478 --MIKKNPASP--HIKEMKNATLADSSHSIQNSPKLTPKKKKKRRASRKKKEVYRQ 532
478 --MIKKNPASP--HIKEMKNATLADSSHSIQNSPKLTPKKKKKRRASRKKKEVYRQ 532
Db 627 PDSIERRPVVDGGGTNSITVYRNATFTW----- 653
627 PDSIERRPVVDGGGTNSITVYRNATFTW----- 653
QY 533 LQREHQAVALAEQGHLLDSDERSPEEBEGKHHGLRLQRTLSIDLEIOEKGVLG 592
533 LQREHQAVALAEQGHLLDSDERSPEEBEGKHHGLRLQRTLSIDLEIOEKGVLG 592
Db 654 -----ARSDPP-----TLNGITPFSIPEGALVA 675
654 -----ARSDPP-----TLNGITPFSIPEGALVA 675
QY 593 ICGSVSGKTSLSAIGOMTLLEGSLAISGTFAVVAQAMINATLRDILFEGEKDEE 652
593 ICGSVSGKTSLSAIGOMTLLEGSLAISGTFAVVAQAMINATLRDILFEGEKDEE 652
Db 676 VVGQVGGGKSLSLALLAEMDKVGHVAKGSAVYVQAQMIQDSLRNLEFGCQLEEP 735
676 VVGQVGGGKSLSLALLAEMDKVGHVAKGSAVYVQAQMIQDSLRNLEFGCQLEEP 735
QY 653 RYNSVLSGCCIRPDLALPSSDLTEIGERGANLSGGORRISLARALYSDRSYIILDDPL 712
653 RYNSVLSGCCIRPDLALPSSDLTEIGERGANLSGGORRISLARALYSDRSYIILDDPL 712
Db 736 YRSVYQACALLPDLTEIPSGDRTEIGEGVNLGGOKORVSLARAYASNADYILDDPL 795
736 YRSVYQACALLPDLTEIPSGDRTEIGEGVNLGGOKORVSLARAYASNADYILDDPL 795
QY 713 SALDAHNGHIFNSAI--RKHLSKTVLFTVHOLOYVLDDEVIFMEGCTTERGTHEEL 770
713 SALDAHNGHIFNSAI--RKHLSKTVLFTVHOLOYVLDDEVIFMEGCTTERGTHEEL 770
Db 796 SAVDAHVGKHLFENVIPKGMKLNKRLILVTHSMSTYLPQVDVILVMSGKRISEKGIQEL 855
796 SAVDAHVGKHLFENVIPKGMKLNKRLILVTHSMSTYLPQVDVILVMSGKRISEKGIQEL 855
QY 771 MNLNGDYATIF-----NNLL-----GETPVEIN 795
771 MNLNGDYATIF-----NNLL-----GETPVEIN 795
Db 856 LARDGAFAEFLRTVSTEQBDADENGVTGVSQPKAKQENGMVLVTSAGKQLQKOLS 915
856 LARDGAFAEFLRTVSTEQBDADENGVTGVSQPKAKQENGMVLVTSAGKQLQKOLS 915
QY 796 SKKETSQOKKSQDKGPKRTGSIKKEKAVKPEEGOLVLEEKGGGSVMSYGYVIAQAG 855
796 SKKETSQOKKSQDKGPKRTGSIKKEKAVKPEEGOLVLEEKGGGSVMSYGYVIAQAG 855
Db 916 SSSYSQSDISRHNN--STAELQKAEKKEETWIKMEADRAQOTQOVVLYSWMDTKAIGL 972
916 SSSYSQSDISRHNN--STAELQKAEKKEETWIKMEADRAQOTQOVVLYSWMDTKAIGL 972
QY 856 PLATVIMALFPLVGSSTASTWMLSYWIKQSGNTVYTGNETSVSDSKNDPHMGVYA 915
856 PLATVIMALFPLVGSSTASTWMLSYWIKQSGNTVYTGNETSVSDSKNDPHMGVYA 915
Db 973 FISFLSIF-LFMCNHSVSLASNTWLSLMTDPIVNGT--QEHKVRSLVYAGALISQGI 1028
973 FISFLSIF-LFMCNHSVSLASNTWLSLMTDPIVNGT--QEHKVRSLVYAGALISQGI 1028
QY 916 SIYALSMVMLILKAIGVVFVKGLTRASSRLHDELFRRLRSLPMKFFEDTPTGRIILNR 975
916 SIYALSMVMLILKAIGVVFVKGLTRASSRLHDELFRRLRSLPMKFFEDTPTGRIILNR 975
Db 1029 AVFGYSMAVSI-----GGLASRCLHVDLHLSILRSFMSFFERRPSGNCILNRF 1076
1029 AVFGYSMAVSI-----GGLASRCLHVDLHLSILRSFMSFFERRPSGNCILNRF 1076
QY 976 SKMDDEVVRLPQAEKFIQNVILVFCVGMIAGVFPWFLVAVGPLYILSVLHYSRVL 1035
976 SKMDDEVVRLPQAEKFIQNVILVFCVGMIAGVFPWFLVAVGPLYILSVLHYSRVL 1035
Db 1077 SKELDVTDSMIEVIRKFMGSLFENVIGACIYIILATPIAIIIPPLGLI---FFVQRFY 1133
1077 SKELDVTDSMIEVIRKFMGSLFENVIGACIYIILATPIAIIIPPLGLI---FFVQRFY 1133
QY 1036 I---RELKRLDNTQSPFLSIITSSIOGLATIHAYNKGQERLHRYQELLDNDQAPFLFT 1092
1036 I---RELKRLDNTQSPFLSIITSSIOGLATIHAYNKGQERLHRYQELLDNDQAPFLFT 1092
Db 1134 VASSRQRLKLESVSRSPYSHFNETLLGVSYIRAFEEQERIRHOSDLKVDENQAVYPSI 1193
1134 VASSRQRLKLESVSRSPYSHFNETLLGVSYIRAFEEQERIRHOSDLKVDENQAVYPSI 1193
QY 1093 CAMRLAVRLDLISIALITTTGLMIVLMHGOIPRAYAGLAISYAVOLTLGFOFTVRLASE 1152
1093 CAMRLAVRLDLISIALITTTGLMIVLMHGOIPRAYAGLAISYAVOLTLGFOFTVRLASE 1152
Db 1194 VANRWLAVERLECVGNCIYLAALFAVISRHSLSAGVLGSLVSYLOVTTYLLNMLVRRSSE 1253
1194 VANRWLAVERLECVGNCIYLAALFAVISRHSLSAGVLGSLVSYLOVTTYLLNMLVRRSSE 1253
QY 1153 TEARFTVERINHYIKTSLAPARINKKAPSPDQEGEVTEBNAMRRRENILPVLYKK 1212
1153 TEARFTVERINHYIKTSLAPARINKKAPSPDQEGEVTEBNAMRRRENILPVLYKK 1212
Db 1254 METNIVAVERLKEYSET-EKEAPWQIOETRPSPSPQVGRVEFRNYCYLRYREDLDFVLRH 1312
1254 METNIVAVERLKEYSET-EKEAPWQIOETRPSPSPQVGRVEFRNYCYLRYREDLDFVLRH 1312
QY 1213 VSPFIKPKREKIGYGRGSGSSIGMALFRLVELSGGCIKIDGVRISDIGLADRSLKSI 1272
1213 VSPFIKPKREKIGYGRGSGSSIGMALFRLVELSGGCIKIDGVRISDIGLADRSLKSI 1272

Db 1313 INVINGEKVQIVGRTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKITI 1372
QY 1273 IPOEVLGSGTVRSNLDPEFNOYTEDQIMDALEPRTMKECIAQLPKLSEVWENGDSNV 1332
Db 1373 IPDQVLTSGSRMLDLDPFSQSDSEWVTSLELAHKDFVSALPKLHCEGEGENLSV 1432
QY 1333 GEROLICARALIRKIKLIIIDEAFAADTEFDLLIOETIRAFADCTMLTAHRLHYVL 1392
Db 1433 GQROLVCLARALIRKIKLIIIDEAFAADTEFDLLIOETIRAFADCTMLTAHRLHYVL 1492
QY 1393 GSDRIMVLAQGOVEFDPSPVLSLNDSSRFYAM 1425
Db 1493 DTRVIVLDKGEIOEYGAUSDLL-QORGLFYSM 1524
RESULT 14
US-08-461-384B-2
Sequence 2, Application US/08461384B
Patent No. 6025473
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,384B
FILING DATE: 05-JUN-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01547
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-384B-2

Query Match 28.4%; Score 2075; DB 3; Length 1531;
Best Local Similarity 34.5%; Pred. No. 9e-195;
Matches 487; Conservative 261; Mismatches 481; Indels 184; Gaps 21;

QY 100 PVDNAGLSCMTFSWLSLARIARAKKGEISMEDVWSLSKHSSDYNCRLERLMOBEELNE 159
Db 209 PSSASFLSRITRFEWITGLI-VRGYRQPLEGSDLSLNKEDTSEQVVPVLVNMWKECAK 267
QY 160 V-----GPDAA-----SLRRVWVIFCRRLTL 181

Db 268 TRKQPVKVVYSSKDDAOPKRESSKVDANEVEALIVKSPQKEMNPSPFLVLYKTFPGPYFLM 327
QY 182 STVCLMTIQLAGFSGPAPFVWKHLEYTOATESNLOYSLILVGLLTELVEKWSALATYA 241
Db 338 SFFFAFIHDLMMFSGPO-ILKLLIKFVNDTKAPDMQGYFYVLYLVLTACIQLTVLHQYFH 386
QY 242 LMYFQVRLRGAILMAFKEKILKINKEKS--LGELINIGSNDGORMEAAVYSLAG 299
Db 387 ICFVSGMRKRTAVICAVYIKKALVITNSARKSSTVEIYVLMVSDQREFDLATYINMIS 446
QY 300 GPVVAILGNIYVNIILGPTGFLGSAVFILFYPAAMFASRLTAFFRKCAVATDERVOKAN 359
Db 447 APLOYIALYILMLNMGPSVLAVAGVAVMLVPMVNVMAKTKTYOYAHMKSKDNRLKILN 506
QY 360 EYLYTIKFKIYAWYKASQSVOKITREERRLLEKAGFQSTITGVAPLVVYIASVYFS 419
Db 507 EILNIGIKVLKTLAMELAEFKVLAIROBELKVKKSAYASAGTFETWVCTPPLVALCTFA 566
QY 420 VHMTLGFD--LTAAGAFVYVYVNSMTFALKVTPPSYKSLSEASVAVDRFKSLFIMEEVH 477
Db 567 VYVTLIDENNIIDAGTAFAVSLALFNILRPPLNLPYVISTIVASVSLKRLFLSHEEL 626
QY 478 --MIRKPPASP---HIKTEKNATLWDSHSSSIONSPKLTPEKMKDKKRAKREKRVQ 532
Db 627 PDSIERRPVKDGGTNSITVRNATFTW----- 653
QY 533 IQRTEHQAVLABOKGHLILLDSDERPSPEEKGKHIHLGLRLQRLHSIDLEIQEGLVG 592
Db 654 -----ARSDP-----TLNGITFSIPEGALVA 675
QY 593 ICGVSGSKTSLISALIGOMTLLGESSIAISGTFAYVAQOAMLTNLTIRNIIIFGKEYDE 652
Db 676 VVGQVCGCGKSLSLALAMDRVEGHVALKGSVAYPAQAMONSLRENNILFGCOLERP 735
QY 653 RYNSVNSCCLRPDLAIIPSPDLTEIGERGAMLSGQROKISLARLYADRSIYIIDLPL 712
Db 736 YRSVYIACALLPDLIELPISGDRTEIGEGKGVNLSGQKRVSLARVYNMAYIYLFDDL 795
QY 713 SALDAVGNHIFNSAI--RKHLKSKTVLFVTHQLOYLVDCDEVIEFMKEGCIERTGTHEEL 770
Db 796 SAVDAHVGHIFENYVIGPGLMKKNKTRILVTHSMSTYLPQVDVIVSGKISEMGSYQEL 855
QY 771 MNLNGDYATIF-----NNULL-----GETPVEIN 795
Db 856 IARDAFAEFLRTYASTEGQDAENGVTGVSQPKGAOKOMENGMLVYDTSACKOLOROLS 915
QY 796 SKKETSGOKSODKPGKSGIKKEKAVPEBGOLVQLEKGGSGVPMVYGVYIOAAG 855
Db 916 SSSSYSGDISRHHN---STAEIQKAEAKKEETWKLAEADKAQTGOVKLSYVDYMKAIQL 972
QY 856 PLATFVIMALEFNLVNGSTAFSTWMLSYWIKOGSGNTVTRGNETSYSMDKNPHQVYA 915
Db 973 FLSPLSLIF-LFMCNHVSAALSNYWLSTMTDDPIVNGT--QEHYKRLVSLGALISQGI 1028
QY 916 STYALSMAYMLILKAIKGVYVYKTRASSRLADELFRILSLSPKAFPTTPTGRILNEF 975
Db 1029 AVFGYSMAVSI-----GGILASRCLHVDLHSILSPMSFFERTSPGMLVNER 1076
QY 976 SKDMDVDVRLPFOAEMFQNIYVLFVFCGMVAGVPMVLVAVPVILFSLVLIHSRYL 1035
Db 1077 SKELDTVDSDMIEPVYIKMEGSLFPNVIACIVLILATPIAIIIPPLGLTY---FFVQRY 1133
QY 1036 I---BELKRLDNIQSPLSHITSSIOGLATIHAVNKGGEFLHRYQELLDNQAPFFLT 1092
Db 1134 VASSSOLKRLKESVSSPYVSHNETLLGVSVIRAFEGORRFTIHSQDLKADQNAKAYPSI 1193
QY 1093 CAMRLAVRLDLSIALITTTGLMAYLMHGOIPRAYAGIAISYAVQULRPFYVRLASE 1152
Db 1194 VANRLAVRLCEVGNACIVLFAALFVISHRSLSAGIAGVSYSLOVYTYLWLVLVMSSE 1253
QY 1153 TEARFYSVERIHYYIKTSLSEAPARIKNAKAPSPDMQOEVEVFEANAEMRYRENPLVLAK 1212


```
OY 1153 TEARTSYERINHVIKTLSEAPARIKKNKAPSPDPPOGEVTFENAEEMRYRENLPVLKK 1212
      | :|||: | :| ||| :| :| ||| | | | :|||:| ||:
Db 1254 METNIVAEERLKEISET-EKEAPMOIOETRPSSNPQVGRVFEFRNYCLRYREDLDFVLRH 1312
OY 1213 VSFTIKPEKIKIGIVRTGSGKSSLGMAFLRVELSGCCIKIDGVRISDIGHADLRKLSI 1272
      :| | | :|||:| | | :| | :| | | :| | | :| | | :| | :| |
Db 1313 INVTINGEKYIGVIGRTGAGSSLTGLFRINESAGEI IIDGINIAKIGLHDLRFKITI 1372
OY 1273 IPOEPVLFSGTVRSNLDPPNOYTEDQIWDALERTHMKECIAQLPKLESEVMENGDNFSV 1332
      |||:| | | | :| | | | :| | :| | :| | :| | :| | :| |
Db 1373 IPODPVLFSGSLRMNLDPPFSQSDSEVWTSLELAHLKDFVSALPDKLDHECAEGEENLSV 1432
OY 1333 GEROLLCIARALLRHCKTLIIDEAFAAMDTEDDLIOETIREAFADCTMLTAHRLHTVL 1392
      | :| | :| | | | | | | :| | | | | | | | | | | | | | | | :| :|
Db 1433 GOROLVCLARALLRKTKILVLEDAFAAVDLETDLLIOSTIRTOFEEDCTVLTIAHRLNTIM 1492
OY 1393 GSDRIMVLAQOVVEFDPSPVILSNDSSRFYAM 1425
      | :| | :| | :| | :| | | | | | | | | | | | | | | |
Db 1493 DYTRVIVLDKGEIQEYGAPSDLL-QQRGLEFYSM 1524
      | :| | :| | :| | :| | | | | | | | | | | | | | | |
```

Search completed: November 9, 2001, 17:46:43
Job time: 6771 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 15:58:52 ; Search time 68.68 Seconds

(without alignments)
1593.809 Million cell updates/sec

Title: US-09-528-031-2

Perfect score: 7308
Sequence: 1 MKDDICKEITIPSPGYRSV.....DSRFYAMFAAEKNAVKG 1437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4811	65.8	946	1	JC5667
2	2371.5	32.5	1427	2	T20903
3	2308	31.6	1400	2	T20904
4	2100.5	28.7	1541	1	S71839
5	2075	28.4	1531	1	DVHDAR
6	2068	28.3	1545	1	S71841
7	2032	27.8	1495	2	E86428
8	2023.5	27.7	1622	2	D86428
9	1992	27.3	1623	2	T01369
10	1983.5	27.1	1515	1	S51863
11	1941	26.6	1527	2	UE0336
12	1928.5	26.4	1488	2	E86428
13	1897	26.0	1573	2	T21219
14	1882	25.8	1539	2	T48059
15	1874.5	25.6	1478	2	T38712
16	1828	25.0	1355	2	T00961
17	1828	25.0	1354	2	T52080
18	1824	25.0	1345	2	T46645
19	1819.5	24.9	1516	2	F84919
20	1811	24.8	1515	2	F52081
21	1782.5	24.4	1545	2	T42751
22	1781	24.4	1546	2	T42728
23	1772.5	24.3	1490	2	T47840
24	1767	24.2	1442	2	T27408
25	1749	23.9	1592	2	S48933
26	1741.5	23.8	1511	2	T42711
27	1733.5	23.7	1559	1	S64757
28	1727	23.6	1477	2	S64616
29	1723	23.6	1389	2	T47796

30	1713.5	23.4	1398	2	T20434	hypothetical prote
31	1669.5	22.8	1582	2	A56248	sulfonylurea recep
32	1665	22.8	1502	2	T42216	multidrug resistan
33	1653	22.6	1421	2	T34225	hypothetical prote
34	1649.5	22.6	1631	2	S64800	probable membrane
35	1597.5	21.9	1037	2	T50518	ABC transporter-11
36	1521	20.8	1548	1	DVLNS	multidrug resistan
37	1464	20.0	1146	2	F84487	probable ABC trans
38	1454.5	19.9	1724	2	T18343	P-glycoprotein - S
39	1451	19.9	1492	2	A39322	cystic fibrosis tr
40	1450	19.8	1451	2	T30821	P-glycoprotein - S
41	1417.5	19.4	1677	2	T18344	P-glycoprotein E -
42	1408	19.3	1480	1	DVHCF	cystic fibrosis tr
43	1393	19.1	1476	1	A40303	cystic fibrosis tr
44	1386.5	19.0	1469	2	T50210	probable ABC trans
45	1382	18.9	1153	2	T26883	hypothetical prote

ALIGNMENTS

```
RESULT
1
JC5667
multidrug resistance protein, short type - human
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C:Accession: JC5667
R:Suzuki, T.; Nishio, K.; Sasaki, H.; Kurokawa, H.; Saito-Obara, F.; Ikeuchi, T.; Tan
Biochem. Biophys. Res. Commun. 238, 790-794, 1997
A:Title: cDNA cloning of a short type of multidrug resistance protein homologue, SMRP
A:Reference number: JC5667, M0ID:97472289
A:Accession: JC5667
A:Molecule type: mRNA
A:Residues: 1-946 <SUZ>
A:Cross-references: DDBJ:AB005659; NID:q2554609; PIDN:BAA22887.1; PID:q2554610
A:Gene: GDB:ABCC5; MRP5; SMRP; ABC33; MOAT-C
A:Cross-references: GDB:9954943
A:Map position: 3q25-3q27
C:Superfamily: human multidrug resistance protein, short type; ATP-binding cassette h
C:Keywords: ATP; nucleotide binding; P-loop
F:87-268/Domain: ATP-binding cassette homology <ABCL>
F:104-111/Region: walker A motif
F:104-111/Region: nucleotide-binding motif A (P-loop)
F:214-219/Region: walker B motif
F:719-912/Domain: ATP-binding cassette homology <ABC2>
F:736-743/Region: nucleotide-binding motif A (P-loop)
F:736-743/Region: walker A motif
F:859-863/Region: walker B motif

Query Match      65.8%; Score 4811; DB 1; Length 946;
Best local Similarity 99.8%; Pred. No. 1.3e-315;
Matches 944; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 492 MKNATLAMDSSSHSSIONSPLTPKMKDKRASGKKEKVRQLOPTEHQAVLADQKGLL 551
      |||||||
DB 1 MKNATLAMDSSSHSSIONSPLTPKMKDKRASGKKEKVRQLOPTEHQAVLADQKGLL 60

OY 552 DSDERPSPREEEGKHHLGLRLORTLHSTDLFOEKKIVGIGSVSGKTSLSALIGQ 611
      |||||||
DB 61 DSDERPSPREEEGKHHLGLRLORTLHSTDLFOEKKIVGIGSVSGKTSLSALIGQ 120

OY 612 MTLLEGSIAISGTFAYVAQAWILNATLRDNIIFGKEYDERVNSVLSGCLRPDLAIIIP 671
      |||||||
DB 121 MTLLEGSIAISGTFAYVAQAWILNATLRDNIIFGKEYDERVNSVLSGCLRPDLAIIIP 180

OY 672 SSPLTTEIGERGAMLSGGQORISLARALYSRSTIYIIDPLSLADAHVGHIFNSAIRKH 731
      |||||||
DB 181 SSPLTTEIGERGAMLSGGQORISLARALYSRSTIYIIDPLSLADAHVGHIFNSAIRKH 240

OY 732 LKSKTVLFTVHLOLYVDCEVIFMKEGCTTERTGETHELNINNGDVATITNNLLIGETPP 791
      |||||||
```

Db 241 LKSTVLFVTHQLOYLDCDEVIFMKECITERGHEELMNLNGDIATIFNNLLIGETPP 300
 QY 792 VEINSKRETSQSKSODKPKRTGSIKKEKAVKPEEGQVLOLEEKGGSPVSYGVYIQ 851
 Db 301 VEINSKRETSQSKSODKPKRTGSIKKEKAVKPEEGQVLOLEEKGGSPVSYGVYIQ 360
 QY 852 AAGGRLATLVLMALFMLNVGSAFSTWMLSVWIKOGSGNTVTNRGNENSVSDSMKDNPHM 911
 Db 361 AAGGRLATLVLMALFMLNVGSAFSTWMLSVWIKOGSGNTVTNRGNENSVSDSMKDNPHM 420
 QY 912 QYASIVYLSMAVMILKAIGVVFVKGTLRASRLHDELFRRIILSRPMKFFDTPTRI 971
 Db 421 QYASIVYLSMAVMILKAIGVVFVKGTLRASRLHDELFRRIILSRPMKFFDTPTRI 480
 QY 972 LNRFSKDMDEVDRPLPFOAEMFIQNVILVFCVGMAGVFPWFLVAVGPLVILFSLHIV 1031
 Db 481 LNRFSKDMDEVDRPLPFOAEMFIQNVILVFCVGMAGVFPWFLVAVGPLVILFSLHIV 540
 QY 1032 SRVLIRELKRLDNTQSPFLSHITSSIOGLATIHAYNKGQELHRYQBELDNOAPFELF 1091
 Db 541 SRVLIRELKRLDNTQSPFLSHITSSIOGLATIHAYNKGQELHRYQBELDNOAPFELF 600
 QY 1092 TCAMRWLAVRLDLISALITTTGLMIVLMHGOIPPAYAGLAISYAVOLTGLFQFTVRLAS 1151
 Db 601 TCAMRWLAVRLDLISALITTTGLMIVLMHGOIPPAYAGLAISYAVOLTGLFQFTVRLAS 660
 QY 1152 ETEAFSTVERINHYIKTSLSEAPARIKNAKSPDMPQEGEVTEFENAEKRYRENTPLVVK 1211
 Db 661 ETEAFSTVERINHYIKTSLSEAPARIKNAKSPDMPQEGEVTEFENAEKRYRENTPLVVK 720
 QY 1212 KVSFTIKRKEKIGIVGRSGSKSLGMALFLRVELSGGCIKIDGVRISDIGIADLRSKLS 1271
 Db 721 KVSFTIKRKEKIGIVGRSGSKSLGMALFLRVELSGGCIKIDGVRISDIGIADLRSKLS 780
 QY 1272 IIPQEPVLFSGTVRSNLDLPFNOYTEDQIMDALERTHMEKCIQAOLPLKLESEVMENGDNFS 1331
 Db 781 IIPQEPVLFSGTVRSNLDLPFNOYTEDQIMDALERTHMEKCIQAOLPLKLESEVMENGDNFS 840
 QY 1332 VGEROLLCIARLRLHCKKILIDEXTAAMDPTDILLIOETREARADCTMLTIARHLTV 1391
 Db 841 VGEROLLCIARLRLHCKKILIDEXTAAMDPTDILLIOETREARADCTMLTIARHLTV 900
 QY 1392 LGSDDIMVLAOGVVEFDPSPVILSNDSSRFYAFMAAENKVAHVKG 1437
 Db 901 LGSDDIMVLAOGVVEFDPSPVILSNDSSRFYAFMAAENKVAHVKG 946
 RESULT 2
 T20903
 hypothetical protein F14F4.3b - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T20903; T22217
 R:Wallis, J.
 submitted to the EMBL Data Library, January 1998
 A:Reference number: Z19344
 A:Accession: T20903
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1427 <W11>
 A:Cross-references: EMBL:AL021446; PIDN:CAB54225.1; GSPDB:GN00028; CESP:F14F4.3b
 A:Experimental source: clone F14F4
 R:Petcy, C.
 submitted to the EMBL Data Library, December 1996
 A:Reference number: Z19531
 A:Accession: T22217
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1427 <W12>
 A:Cross-references: EMBL:Z83227; PIDN:CAB54250.1; GSPDB:GN00028; CESP:F14F4.3b
 A:Experimental source: clone F45B8
 C:Genetics:
 A:Gene: CESP:F14F4.3b

A:Map position: X
 A:Introns: 26/2; 61/2; 116/2; 140/2; 218/2; 306/3; 356/1; 478/1; 534/3; 642/3; 669/3;
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homolo
 Query Match 32.5%; Score 2371.5; DB 2; Length 1427;
 Best Local Similarity 36.3%; Pred. No. 4,66-151;
 Matches 507; Conservative 289; Mismatches 497; Indels 103; Gaps 20;
 51 ETARAREGISLSDASMSHQLRIIDEEHPKRYHHGLSALKPIRTCKHOH----PYDNAGL 106
 26 ETSDRAKRYAGDTNKT-----IGRYSAVQNLPIPTTERNNKNGSRIDDAGL 75
 107 FSCMTFSLSLARVAHKKGELSMEDVMSLSKHSSEVDNCRLELMQELNEVPDAS 166
 76 FSEFTVSWVFPYLYQA-VRGKLDNRQVWGCSFSDSGINMAALEVLEWDE-KKANAKSPS 133
 167 LRRVWIFCRFRLISIVCLMITOLAGFSGPAFWKHLELYQATSENIQO-----YSL 220
 134 LFKVYIRFISTRLMFSCAVFFECILFGFIQPCFIRLLIAFENREDOBSRIYVYGLA 193
 221 LVGLLLEIYRSMSLATWALNFRGVRLRGAILTMAFKLIKLIKERSLGEELNIC 280
 194 LVAAISVEEFARVLISYATMAVSYRTGIRVGAVLALLYKNVLNSKDLGCTESDVINIF 253
 281 SNDGGRMEAAVGSLLAGRPVALIGMIVYIILGPGFLGSAFIILFYRAMMFASRLT 340
 254 ANDGRLDVAFTAPLVALPLVGLVGGIGYLLMIVIGRMSLGILFEVFDVYIQGLGSM 313
 341 AFRRRKCVAAEDERYOKMNEVLTLYIKFMAYVAFSOSYOKIREBRRLIEKAGYOS 400
 314 VACRLALVYKTEKRISMAEIIKIRIVKMNQWEDIFSAKIDQFKEREKVOGRKGYAOS 373
 401 ITVGAPIVYVIVASVYTSVHMTLGFDLTAQAFIVYVFNMTALVYTPSVKSLSEA 460
 374 LAIAGCPVVPVVAALLTFVGVALNDLASDAFSAIYVFMELGIRIIPGSRYLEA 433
 461 SVAVDRFKSLFMEEV--HMINKPASPHIKEMKNATLAMSSSHSSIONSFKLPKKMK 518
 434 VYAMRRIQETYLLEOAYRPPVYN---AEDVULDCGATITTYO-----PK--- 474
 519 DKRASGRKKEVROLQRTENQAVLAEOKGHLLDSDERSPEEBGKHIGHRLQRTL 578
 475 ---AAKAPVDETK--EPENEVIYVETPVF-----TC 501
 579 HSIDLEIDQKLVGICGVSQKTSLSAIILOQMTLSESTAIS-GTPAYVAQAWIINA 637
 502 SPDKLSIKRGEHIAVIGAVGCGKSAIILKISGMFTTDALSDRSQTVYVYOKAMIFNG 561
 638 TLRDNLIFGKEYDERRYNVLNSCCIRPLALIPSSDLTEIGERGANLSGGORISLAR 697
 562 TVQDNILIFGDKKNSERYIYKAVNGCQLTEDLTLLSVDGRTEVEGERKATILSGGKAVALAR 621
 698 ALYSRSRIYILDPLSALDAHYGNHIFNSAIRKHLKSKTVLFVTHQLOYLDCDEVIFMK 757
 622 AVFGTKNLYLFDIDFASLDKVAANKIHEEIIQKLKKKALMVTNNMELHHFDRVLVE 681
 758 ECGITERGTHEELMNLNGDIATIFNNLLIGETPPVPINSKKTSSQSKSODKPKRTGSI 817
 682 GGNIVADGNHDLIERKNDAKYKTFVD-----ACEYQATSGATSPCGGPAQAPL 731
 818 KKE-----KAYKPEEGQVLOLEE-KGGSVPMYVGYVYIOAAGPLAFLVLMALFMLNV 870
 732 DAEIILNNSSEDLKGDADKLISDEEDMGNSTIMWRIYKOYIHAAAGMPITVTCIGFIYAV 791
 871 GSTAFSTWMLSVWIKOG-SGNTVTNRGNE-TSVSDSMKDNPHMQYASIVYLSMAVMIL 928
 792 VSNISTVYLSRWLKKKHGHETTITNGTEFLMKTSIADSPTGYAAVYVALVLTIS 851
 929 KAIRGVVFKGTLRASSRLHDELFRRIILSRPMKFFDTPTRIILNRSKDMDEVDRPLF 988
 852 GLFKACVFKVSLTAATRLHDMFOAVIHGATSFDSPTPTGRIILNRSKDMDEVDRPLF 911

Oy	989	CAEMTQVAVILFECVG---MIAGYFPFLVAVGPVLIVLESYLHVSAYLIRELRDNI	1045
Db	912	TAEVFLLQMVI--CLGLGVITSVFYPLFLFAIPLFVFEVSCFACIRNLRSEHI	968
Oy	1046	TOSPLSHITSISIOGIATIHAYNKGOEFIAHRYOELLDNOAPFFLETCAMRIARVDLI	1105
Db	969	SRSPIYDHVASLECITIITHFFOQSNEPFLVKRLHDNSGALIFEFOSAKRWLAWDLI	1028
Oy	1106	SIALTTTGAMIIVMHGIOIPPAVAGIAISAVALGTGLEPFTVRSLASEAFERTSVERINH	1165
Db	1029	VVMVAIIAALLTVMLTGVSPADAGMAIAFAVQMSGICGFVANRQTETEAMTSERASY	1088
Oy	1166	YIKTLISLEAPARKKAP-SPDWPEOGEEVTENEMARMREMPLYLKRVSTTIKPKKIG	1224
Db	1089	IADNIPEEGEWNTRGGDIIESSWPANGINFESVULRYRKSHPIALANDITEIKGEKEVG	1148
Oy	1225	IVGRSGSKSSLGMALFPLVELSGGCIRKDGVRTSDIGLADRSLSTIIDPEPVFSGETV	1284
Db	1149	IIGRGSSKSSLANLIFLPYPTWTGTYIIDGDVDRTELVKLRGISAIADDPSESSTV	1208
Oy	1285	RSNLDPENQYTEODIMDALERTHMKECIAQLPKLSESEVENGDNFYSGEROLCTIARAL	1344
Db	1209	RPNIDPSLEYSDSMIWELAERCHKTIVQSLDKLEADVSHGNNGSFSGEQRLFARAL	1268
Oy	1345	LRRCKILLIDATAAMDTEDLLIOETIREAPADCTMTLTIAHLHTVYGSPRIYVAOQG	1404
Db	1269	LKRSRITVIDLATASVPDAQDKLIDOEVIKTFVADATVIIAHRLDNVNMDRMHLMKKG	1328
Oy	1405	VVEPDTPSVLSNDSS 1420	
Db	1329	LINFTTPQEKMRDMS 1344	
RESULT	3		
	T20904		
	hypothetical protein F14P4_3a - Caenorhabditis elegans		
C:	\$Species: Caenorhabditis elegans		
C:	Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000		
C:	Accession: T20904; T22218		
	R:Wallis, J.		
	submitted to the EMBL Data Library, January 1998		
A:	Reference number: Z19344		
A:	Accession: T20904		
A:	Status: preliminary; translated from GB/EMBL/DDBJ		
A:	Molecule type: DNA		
A:	Residues: 1-1400 <M12>		
A:	Cross-references: EMBL:AL021446; PIDN:CAB54226.1; GSPDB:GN00028; CESP:F14P4_3a		
A:	Experimental source: clone F14P4		
	R:Percy, C.		
	submitted to the EMBL Data Library, December 1996		
A:	Reference number: Z19531		
A:	Accession: T22218		
A:	Status: preliminary; translated from GR/EMBL/DDBJ		
A:	Molecule type: DNA		
A:	Residues: 1-1400 <M12>		
A:	Cross-references: EMBL:T83227; PIDN:CAB54251.1; GSPDB:GN00028; CESP:F14P4_3a		
A:	Experimental source: clone F45B8		
	C:Genetics:		
	A:Gene: CESP:F14P4_3a		
	A:Map position: X		
	A:Introns: 26/2; 61/2; 116/2; 140/2; 218/2; 306/3; 356/1; 478/1; 534/3; 642/3; 741/2; 767/2		
	C:Superfamily: human multidrug resistance protein cMOA2; ATP-binding cassette homology		
	Query Match	31.6%; Score 2308; DB 2; Length 1400;	
	Best Local Similarity	35.6%; Pred. No. 8.4e-147;	
	Matches 497; Conservative 285; Mismatches 484; Indels 130; Gaps 21;		
Oy	51	ETARAEGSLDASHNSOLRIDDEHPKGYHHGSALKPIPTCKHOH----PYDNAGL 106	
Db	26	ETSDAKRYAGDTNRKT-----IGRYSAVQNLIKPLRTTKRNKNKGSRIDDAGL 75	
Oy	107	FSCMTFWSLSLARYAAHKKGLSMEDVWSLSKHESDVNCRLERIOMDELINENVGDAAS 166	

Dd	76	FSEFYTSMVEPPYLYQ	A	V	R	K	E	L	D	R	N	O	W	C	S	F	D	S	C	G	L	M	A	R	E	V	L	M	D	E	-	K	A	N	A	K	S	P	S	133
Qy	167	LRRVWVJFCRTRLISIVCLMITQI	A	G	S	G	S	G	P	A	F	M	V	M	C	H	L	E	T	Y	Q	A	T	E	S	N	D	Q	-----	Y	S	L	L	220						
Dd	134	LFKIYFIISIRLMFSCAVFPC	L	I	G	F	I	G	F	I	G	F	I	G	F	I	G	F	I	G	F	I	G	F	I	G	F	I	G	F	I	G	F	I	G	F	I	G	193	
Qy	221	LVLGLLLEIVRSWLSALTALWALN	A	P	T	G	V	R	L	R	G	A	I	L	L	M	A	R	K	I	L	K	I	N	E	S	L	G	E	L	I	N	I	C	280					
Dd	281	SNDQGRFEPAANGSLAGRPVALL	G	M	I	T	A	N	I	L	G	M	I	T	A	N	I	L	G	M	I	T	A	N	I	L	G	M	I	T	A	N	I	L	G	M	I	340		
Qy	254	ANDQGRFEPAALVPLVGRPLV	G	I	G	I	G	I	G	I	G	I	G	I	G	I	G	I	G	I	G	I	G	I	G	I	G	I	G	I	G	I	G	I	G	313				
Dd	341	AYFRKCAVATDERVOYKMNVL	T	I	F	I	K	M	V	A	N	V	K	F	S	O	V	O	K	I	E	E	E	R	L	E	R	K	A	G	F	O	S	400						
Qy	314	VACRNLAIVTEKRTISMAELIT	I	T	I	K	V	M	K	N	G	H	E	O	J	F	S	K	I	D	O	F	R	K	E	E	V	O	I	R	K	S	G	I	A	O	S	373		
Dd	374	LAIACGPVPVPAAILTFGVGL	V	L	A	G	N	D	L	L	A	S	D	F	A	I	T	V	E	P	M	L	G	I	R	I	P	K	S	R	Y	L	A	E	A	433				
Qy	461	SVADREKSLFIMEV--	N	H	I	K	K	P	A	S	P	H	I	K	T	E	M	K	A	T	L	A	M	O	S	S	H	S	I	O	N	S	P	K	L	P	K	518		
Dd	434	VVARRRIOEYLLLEQVAPRV	T	N	--	A	E	D	V	L	D	C	O	A	T	Y	T	Y	Q	-----	P	K	---	474																
Qy	519	DKRASRGKEKERVOLQRTHE	O	A	V	A	L	A	O	K	G	H	L	L	D	S	E	R	P	E	E	R	K	I	N	H	L	R	L	R	T	L	578							
Dd	475	---AAKARVDET	K	--	E	P	T	E	N	E	V	I	V	E	P	V	---	T	C	501																				
Qy	579	HSIDLEIOEGKLVGICGSVSG	T	S	L	S	A	L	I	G	M	T	L	E	G	S	I	A	I	S	--	G	T	P	A	Y	A	O	A	M	I	N	A	637						
Dd	502	SFDKLSIKRGHIAVIAVG	A	C	G	S	A	I	K	A	I	G	S	H	M	F	T	T	D	L	A	S	D	R	S	O	Q	T	Y	V	P	O	K	A	M	I	N	G	561	
Qy	638	TLRNIILFGEYDERNNSVLS	C	C	L	R	P	D	L	I	P	S	S	D	L	E	I	E	R	A	N	I	S	G	G	O	R	O	I	S	L	A	R	697						
Dd	562	TVOONILFEGDKMNSERYK	A	N	C	O	L	T	E	D	I	T	L	T	S	V	D	R	T	E	V	E	R	G	A	T	V	L	S	G	K	A	V	A	L	A	R	621		
Qy	698	ALYSDRSYIILDDPLSLAD	A	H	V	M	C	I	F	N	S	A	I	R	K	H	L	S	K	T	V	L	E	V	T	H	O	L	Y	V	D	C	E	V	I	E	M	K	757	
Dd	622	AVFOQKNLYEFDLFA	L	D	K	L	L	H	N	-----	D	V	L	E	654																									
Qy	758	EGCITERGTHEELMNLNG	D	V	A	T	I	F	N	L	L	G	E	R	P	P	E	I	N	S	K	E	T	S	G	O	K	K	S	O	D	K	G	P	R	T	S	I	817	
Dd	655	GNIVADGNHDI	L	E	K	N	D	A	K	T	E	V	D	-----	A	C	E	T	O	A																				

Db 1062 YADNIPEDGEMNTQOGDIIESWPNQINFSEVNLKRSKPLANDITEIKGEKVG 1121
 QY 1225 IIVRTGSGKSSGLMFLRYELSGGCIKIDGVRISDIGLADLRSLIIPQEPVLFSGTV 1284
 Db 1122 IIGRTGSGKSSGLMFLRYELRYVPTNGTITIDGVDIRVGLVLRGISAIADPSLFSGTV 1181
 QY 1285 RSNLDPPNQVTEEDIDWALTEYHKECIAOLPLKLESSEVNGNFGSEGOQLLCTARAL 1344
 Db 1182 RENDPSPLESYSDSMWELECHLTKTQVOSLDKLEADVSHGNNFSGVEKQLEFLARAL 1241
 QY 1345 LRHCIIILDEATAMPTETDILLIQTETIREAFADCTMLTILHRLHTYVLSGRIMVLAQO 1404
 Db 1242 LMKRVIYLDGATRSVADGDKLQEVITKYFADATVYITIIHRLDNRANMRIMHLKNGK 1301
 QY 1405 VVEEDPSPVLLSNDSS 1420
 Db 1302 LINFPTPOEMFKDDMS 1317

RESULT 4
 571839
 canalicular multidrug resistance protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
 C:Accession: S71839
 R:Buechler, M.; Koenig, J.; Brom, M.; Kartenbeck, J.; Spring, H.; Horie, T.; Keppler, D.
 J. Biol. Chem. 271, 15091-15098, 1996
 A>Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance
 A:Reference number: S71839; MUID:96279006
 A:Accession: S71839
 A>Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-1541 <RUE>
 A:Cross-references: EMBL:X6393; NID:91292881; PIDN:CAA65257.1; PID:G1617207
 C:Superfamily: human multidrug resistance protein CMOAT2; ATP-binding cassette homology
 C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein
 F:100-124/Domain: transmembrane #status predicted <TM01>
 F:127-151/Domain: transmembrane #status predicted <TM02>
 F:160-187/Domain: transmembrane #status predicted <TM03>
 F:305-329/Domain: transmembrane #status predicted <TM04>
 F:354-381/Domain: transmembrane #status predicted <TM05>
 F:431-451/Domain: transmembrane #status predicted <TM06>
 F:456-476/Domain: transmembrane #status predicted <TM07>
 F:536-564/Domain: transmembrane #status predicted <TM08>
 F:574-602/Domain: transmembrane #status predicted <TM09>
 F:650-633/Domain: ATP-binding cassette homology <ABC1>
 F:966-994/Region: nucleotide-binding motif A (P-loop)
 F:1018-1046/Domain: transmembrane #status predicted <TM10>
 F:1104-1132/Domain: transmembrane #status predicted <TM11>
 F:1203-1228/Domain: transmembrane #status predicted <TM12>
 F:1313-1506/Domain: ATP-binding cassette homology <ABC2>
 F:1330-1337/Region: nucleotide-binding motif A (P-loop)
 F:6,1007,1010,1011/binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 28.7%; Score 2100.5; DB 1; Length 1541;
 Best Local Similarity 34.4%; Pred. No. 9,1e-133;
 Matches 497; Conservative 262; Mismatches 476; Indels 209; Gaps 26;

QY 98 OHPDNAGLSCMFESWLSLARVAHKKGLSMDVMSL-----SKHSS----- 142
 Db 191 QTSVTSFSSITFSWDTVLKGYKH-PLTLEVDVIDEGCFTRSVTSKFEAMTKDL 249
 QY 143 ----DVCNRLLR-----LWQELINEVGPDAASLRVYVIF-- 174
 Db 250 QKARQARQRLQKQRRPEATLHGLNKKQSQSOQVYLVEKAKKSEKTKRPPKSWMLKS 309
 QY 175 ----CRRLLIISYLCMTQOLAGSGPAFWVKHLLLEYQTATSNLYQSLLVIGILLTEI 230
 Db 310 LKFTFHVYIILKSFILKLIHDLVFLNPO-LIKLILIGVKSNSYVMGYCAIIMFAVTL 368
 QY 231 VRMSLSLALTMALNRTGVRLGALITMAFKKILTKNI--KEKSLGELINICSDGQRMF 288

Db 369 IQSFLQSYFOHCYVLGKCVFTYWSSTYTKKALLTSNLRKQYITIGETVNLMSVDSQKLM 428
 QY 289 EAAVAGSLLAGPVVALIGMTYVNIILGPTGLGSANVILFYPAMMFASRLTAVERKCV 348
 Db 429 DATNMOLVMSVQIOTLSTIFFMRELQPSILAGVAVVLLIPNGVLAFTIRINIQVM 488
 QY 349 AATDERQAKNEVLTLYKFTKMAVAKFASQVQKIREERRIIEKAGYFOSITGVAPI 408
 Db 489 KKKDKRLKIMNEIISGLIKLYFAMWEPFOBOGIRKEKELNLLRFQOISLIFILQI 548
 QY 409 VVVLASVTFESVHTIGFD--LTAQAFTVTVFNMSFALKVTPESVKSSEASVAVDR 466
 Db 549 TPLVSVTFESVYLVDSANVLEKATSTILRNILFPLSMLEPWTSTSLQASVADR 608
 QY 467 FKSLEFMEVMIKRNKPAAPKIK--IEMKNATLWDSHSSIQNSPKLTPKKKKDRASRG 525
 Db 609 LERYLGDDDLTSAIRRSNPNDAKVKFEASFTMD----- 643
 QY 526 KKEVROLQRTQHOAVLAEOKGHLLDSDERPSPDEEGKHILGHLRLQRTLSIDLEI 585
 Db 644 -----PD-----LEATIODVNDI 657
 QY 586 QEGKINGICSGVSGKTSLSAIIIGQWTLLEGSTAISGFAYVAQOAMILNATLBDNLF 645
 Db 658 KPGOLVAAGVTVSGKSSLSAMLEGEMENVHGTITQSTAYVPOQSVITQGTIKDNL 717
 QY 646 GKEYDEERYNSVLNSCCRPDLALIPSSDLTEIGERGANLSGQORORISLARALYSPRSI 705
 Db 718 GSEYNEKKYQOVLAKACALPLLELIPGDMATEIKENGLNSGQKORSLARAAYQADI 777
 QY 706 YILDDPSALDAHYGNHIFNSAIRKH--LKSRYLVFTVHQLOLYVDCDEVYKKECITE 763
 Db 778 YILDDPSAYVAHYAKHIFNKVYVPGNGLACKTRIFVGHIFLPOYDEIVLVKGTILE 837
 QY 764 RGTHEELNLNGDVAITFNLLIETPVEI----- 794
 Db 838 KGSYRDLDDKKGVFARMKKTVMKHSRGEATVNNDSAEDDDGLPTMEIPEDAASL 897
 QY 795 -----NSKET-SSGOKKQODGP-----KTSIKKEKAVPPEGOLVLEKGGQS 840
 Db 898 AMRENSLRRLTSSSRSSRSRSLKNSLKIKVNNVLAKEKEVEGGKLIKKEFVEFGK 957
 QY 841 VPMVSYGYVIOAGCPALFVYLMALFMLNVSSTAFSTWMLSYLKQSGRNTVTRGETS 900
 Db 958 VKFSIYKYLOAVGM-WSILPILFYGLNNAVFGSNLMSAMTSD--SDNL-----NGTN 1010
 QY 901 VSDMKDNPBMQYASTYALSMVAWLILKAIKRGVVFYVGTL-----RASSRLHDELR 953
 Db 1011 NSSSHRD-----KRIQVFGALGLAOGICLLISTIMSYIACRNASKALHGOLLT 1058
 QY 954 RILRSPMKFEDTPTGRILNRFSSKMDDEVYRLPQAEMLIQNVILVFCVG---MIAG 1009
 Db 1059 NILRAPMFFETPTGRIVNRFSGDISFVDLLP----QTLRSMWMCFFGIAGTLVVICM 1114
 QY 1010 VFPWFYVAGVVLVLFVSLHIVSRVILRELKRLDNITQSPFLSHITSIGLATIHAHYNK 1069
 Db 1115 ATPFAFIIILPILSTISYQVYVATSRQRLRDSVYKSPISHSFSEVNGLPITRAFEH 1174
 QY 1070 GQEFLLHRYOELLDNQAPFEFTCAMRLAVRLDISIALTTTGLMIVLHGQOIPAYA 1129
 Db 1175 QORLANNEKQIDNOKCVSWITSNRMLAIRLELVGNLVYFGCALLEVLYRKLTGTGDV 1234
 QY 1130 GLAISYAVOLGTLFQTVYRLASETEARTSEYRINHYIKTSLSEAPARIKKNAPSPMQ 1189
 Db 1235 GFVLSNALNITQTLNMLYRMSSEAFETNVAVERISEYI--VVENENAP--WYDVKRPADMPR 1292
 QY 1190 EGEVTFENAEKRYRENCLPYLKVYSFTIKPEKIGYIGRTGSGSSGLMFLRYELSGG 1249
 Db 1293 HGEIOFNMYQVRYREPDILVLTGKITCNKSGEKGVVGRGAGKSLTNCIFRLIESAGG 1352
 QY 1250 CIKIDGVRISDIGLADLRSLIIPQEPVLFSGTVRSNLDPPNQVTEEDIDWALTEYHMK 1309

Db 1353 QIIIDGIVASIGLHDLERLITIPDPIILPSCGLRMNLDPPNKYSDEBYWALFLAHR 1412
QY 1310 ECTIADLPKLESEVENDNFVSVEROLCTARALLRCKIILIDEAPAMDTETDLIQ 1369
Db 1413 SFVSGIQLGLLESEVEDGNLSIGORQLCLGRAVLRKSKILVLEDAFAVLEFDSLIQ 1472
QY 1370 ETRIRAFDCMPLTAHLHLTVLSGSDRIMVLAQGOVEDPFSVLLSDNSFFYAMF--A 1427
Db 1473 TTRKEFSQCCTVITTAHLRHLTIMDSKIMVLDNGKIVEGSEPELLSNGS-FYLMAREA 1531
QY 1428 AAEN 1431
Db 1532 GLEN 1535
RESULT 5
DVTUHR
multidrug resistance protein (cell line H69AR) - human
N:Alternate names: multidrug resistance-associated protein (MRP)
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1993 #sequence_revision 05-Dec-1998 #text_change 19-Jan-2001
C:Accession: A44231; A37495
R:Coile, S.P.C.; Bhargava, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almqvist, K.C.;
Science 258, 1650-1654, 1992
A:Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer
A:Reference number: A44231; MUID:93088080
A:Accession: A44231
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 'MAPTRSGTMSRGIPATPSPAFRRSSCGGLVFTSPV', 50-1531 <COL>
A:Cross-references: GB:L05628; MID:g1835658
A:Experimental source: small cell lung carcinoma cell line H69AR
A>Note: sequence extracted from NCBI backbone (NCBI:119851); this sequence has been cor
R:Coile, S.P.C.; Deeley, R.G.
Science 260, 879, 1993
A:Title: Multidrug resistance-associated protein: sequence correction.
A:Reference number: A37495; MUID:93262415
A:Accession: A37495
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-60 <CO2>
A:Cross-references: GB:L05628; MID:g1835658
A>Note: sequence extracted from NCBI backbone (NCBI:131929)
C:Genetics:
A:Gene: GDB:MRP
A:Cross-references: GDB:136335; OMIM:158343
A:Map position: 16p13.1-16p13.1
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C:Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transme
F:661-844/Domain: ATP-binding cassette homology <ABC1>
F:678-685/Region: nucleotide-binding motif A (P-loop)
F:788-792/Region: nucleotide-binding motif B
F:1310-1503/Domain: ATP-binding cassette homology <ABC2>
F:1327-1334/Region: nucleotide-binding motif A (P-loop)
F:1450-1454/Region: nucleotide-binding motif B

Query Match 28.4%; Score 2075; DB 1; Length 1531;
Best Local Similarity 34.5%; Pred. No. 4.7e-131;
Matches 487; Conservative 261; Mismatches 481; Indels 184; Gaps 21;

QY 100 PVDNAGLSCMTFMSLSLARAHHKGLSMDVSLSHSSDVNCRKLERLMDEELNE 159
Db 203 PESSASFSLRITFWITGLI-VRGYRQPLEGSDLSLNKEDTSEOVVPLVNMKCECK 267
QY 160 V-----GPDAA-----SLRRVWIFCRRRLIL 181
Db 266 TKQKQYKVVYSSKDRAPQKESKVDANEVEALYKSPKEMNPISLFXKLVITFGYFLM 327
QY 182 STVCLMTQLAGSPAPFWKHLLEYTQATESNLQYSLLVGLLTLTEYRMSIALTWA 241
Db 328 SFFFKAHIDLMFSGPO-ILKLLIKFVNDTKAPDMQGVYFVLLFVTACTQLTLVHQYFH 386

QY 242 LNRRTGVRGALILMAPKKILKLNKEKS--LGLINICNSDQORPEAAVAGSLAG 299
Db 387 ICEVSGMRKTAIVAGVARKALVTNSARKSSTVEIYMLKSVDAQREMDLVTYINMWS 446
QY 300 GVVVALILGIVVWVILIGPFGISAVFILFYPAAMFASRLTAZFAFRKCAATDERVOKAN 359
Db 447 APLQVITLALYLLMLNGLSVLAGVAVVYLMVAVNVMAMKTTYQVAHMKSDNKLMLN 506
QY 360 EVLTYIKFKIYKAWKASQSVOKIREEBRILKEAGFYPOSTGVAPVIVIASVTFPS 419
Db 507 EILNIGIKVLKYAMELAFKDVILRQELKVLKKSAYLSAVGTFWCTPFLVALCTFA 566
QY 420 VAMTIGFD--LTAAGAFVYVYFNSMTFALKTPPSVYSLSSASAAYDRFKLFLMEEH 477
Db 567 YVYTIENNIIIDAQAFVSLALFNLRPLNLPVYISSIVASVSLKRLIFLSHEELE 626
QY 478 --MIKKPASP---HIKTEKNATLWDSSHSSSIONSPKLTFRKMKDKRASGKREKVR 532
Db 627 PSTIERRPYKDGCGNSTIVRNATFTW----- 653
QY 533 LQRTENQAVLABQKHLILDSDEPSPSEEBEGKHILGLRLQRLHSLDLEIQEGKLVG 592
Db 654 -----ARDDP-----TLNGTFSIPGALVA 675
QY 593 ICGSGSGKTSILSAILGOMTLLEGSIASCTPFAVYAOQAMLTNRNILEFKREYDE 652
Db 676 VVGQVCGCGKLSLSLALMEDRVGEHVAIKGSAAVYPOQAMQNSLRNILLFGQDLEP 735
QY 653 RYNSVLSGCLRPDLALIPSDLTEIGERGAMISGQORISILARLYSDRSITYLDDLP 712
Db 736 YRSVITQACALLPDELIPRSGRTIEGKGVNLSSGQKRVSLARAVSNADYILEFDDLP 795
QY 713 SALDAHVGNHFNISAI--RKLKSKTVLEFVTHQOLYVDCDEYIFKBSGCTERGTHEEL 770
Db 796 SAYDAHVGKHIFENYIGPKMLKNKTRILVTHSMSTYLPQVDVILVMSGKISEMSYQEL 855
QY 771 MNLNDYATIF-----NNLL-----GEPFVPEIN 795
Db 856 IARDGAFAEFLRTYVASTEOEDAENGVTGVSQPGKEAKOMENGMIVDSAGQOLORUS 915
QY 796 SKKETSGSQKSDQKPKTKSIRKKEAVKPEEGOLVLEKSGSGVMSVGYVIOAGG 855
Db 916 SSSSYSGDLSRHHN---STAELOKAAKKEEIKMLKMEADKADQGVKLSVYDMYKAIIL 972
QY 856 PLALVIALFELNVLNGSTAFSTWMLSYWIKGSGNTTVTRGNETSVDSKKNPMQYTA 915
Db 973 FISPLSIF-LFMCNHYSAIASNYMLSLMTDDPIVNGT--OERTYRLSYGALGISQGI 1028
QY 916 STYALSMAYMLIKAIRGVVFKGTLRASRLHDELFRILRSPKKEPPTPTGRILNRF 975
Db 1029 AVFGYSMAVSI-----GGILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRF 1076
QY 976 SKDMDVDVRLPFOAEFIONTVLFEFCGMITAGYFPMVLVAVPLVILFSLVHLTVRYL 1035
Db 1077 SKEDLTVDSMTEBVIKMFSGSLFVNIAGACTIVLLATPIAIIIIPLGLY---EPQORY 1133
QY 1036 I--BELKRLDNIOTSPFLSHITSSIOGIATIHAYNKGQELHRRQELLDNOAPEFLTF 1092
Db 1134 VASSRQLKLESSEVSRSPPYSHNETLLGVSIRAEEDRFRHQSLDKDENQAKYPSI 1193
QY 1093 CAMRVLAVRLDLISALITTTGLMTVLMGQIPRAVAGIAISYAVOULGLOPFTYALASE 1152
Db 1194 VANRWLAVLEECVGNCTIVPALFAVISHSISAGLVLSVLSQVTTYLWMLVAMSSSE 1253
QY 1153 TEARFSEVERINHYIKTSLLEAPARIKNKAPSPDMEQEEVEVFENAKMYRENPLVLK 1212
Db 1254 METNIVAVERLKEYSRT-BKEAPWQIOETRPSSWQVGRVERFNNCLAYRDLDOVLNH 1312
QY 1213 VSFITKPKKIGIVGRTSGKSSLGALFLVELSGGCIKIDGVRLSIDGLADLRSKLSI 1272
Db 1313 INVTLINGGEKVGIVGRTGAKSSLTGLFRINESAGEETIIDINIAKIGLDLRKKITI 1372
QY 1273 IPOEPVLFSGVTRSNLDPPNQUTEDQIMDALERTHKKECTIADLPKLESEVENDGNFVS 1332

Oy 1427 AAENKAVK 1436
 | | : |
 Db 1535 AGIENVNSTR 1544

 RESULT 7
 E86428
 Probable ABC transporter [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: E86428
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso
 Chlu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
 C-A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzilli
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: E86428
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1495 <STO>
 A:Cross-references: GB:AE005172; NID:g1055818; PIDN:AMG28288.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 A:Superfamily: human multidrug resistance protein CMOAT2; ATP-binding cassette homology

Query Match	27.88;	Score 2032;	DB 2;	Length 1495;
Best Local Similarity	34.66;	Pred. No. 3,6e-128;		
Matches 486;	Conservative 240;	Mismatches 483;	Indels 196;	Gaps
QY 96	KHNPVDNAGLFCSCMTFSWLSLSTARVAHKGELSMEDWMSLSKHESSDVNCRRERLMOE 155			
DB 224	EHICPERIASIFSFTYIGWITPLMQLYRK-PIREKYVMQDKDYQIETELIKRFQKQWTE 282			
QY 156	ELNEVGPDASLRRWVIF-----CRTLILSLIVCIMITOLAGFSGPAEVMVHLLEYTO 209			
DB 283	E-----SRRPKPWLRLNNSLSGREFWLAGIFKIGDLSQFQPV-ILSHLLRSMQ 332			
QY 210	ATES---NLQYSLLLVGLLTELTVRWSMLATYALNRYGRVRLRGAILTMARKKIILK 266			
DB 333	EGDPAMVGYVAFLIFGVYTLGVGLCEQYFQNW---RQGFRLSTLVAAIFRHSILRLT 388			
QY 267	NIKEKSL--GELINICSDGORMPEAAVGSLLAGGVVAIILGIVYVILIGPTGFLGSA 324			
DB 389	HEAKRNPFGSGVYTMITTDANALQOISQOLHGLMSAPFRITVSMILLYQOLGYASLFGSL 448			
QY 325	VLLFFPAMFASRLTAYFRKKCYAATDERYOKANEVLTYYIKFIKMYAKAFSOSVOKI 384			
DB 449	ILFLPIQLQTLISKMRKRLREGLOMTDKRGVITNELISSMDPVKCYAMKSPESRQGI 508			
QY 385	REERRIIEKA---GYQSIITVGAPLYVYVYIASVYTSYMTKIGFDLTAAQATVTVEN 441			
DB 509	RNEELSWFRKQQLLSANNSFLNSIPV---TVSSGVVLLGGDTLPRAAFTSLFLA 565			
QY 442	SMFPAKLTPEPSVKSLEASAVAYDRFKSLFMEEBVHIKKKPASPHI-KLEMNATLAMD 500			
DB 566	VLRPLNMLPRLLSQVYVANNASTLORIEELLSEBRILLAQNPRLQPGPAISIKNGYRSD 625			
QY 501	SSHSIIONSPLRLPEKMKDKKRAKSGKKEKYVQLOQTEHQAVLAEQKGHLLDSDERSPE 560			
DB 626	S-----KTY----- 629			
QY 561	EEEGKHILHGLRLQRLHSLDLEIOGKLVYIGCGSGSKTSLISAILQOMLLE-GSI 619			
DB 630	-----KPLTSDINTEILPVGTLVAIVGGEKSTSLISMLDELISHAETTSV 675			
QY 620	AISGFAVAQAMVILNATLRDNLFLPEKYEDEERYNVSLNSCCLRPLALIPSSDLTEIG 679			

Db	676	VIRGSVAATVPOWSWIFENATVENLIFGSDFERSEKWRADITAFALQHDLDLPGKDLEITG	735
Qy	680	ERGAMISGSGORISIIAALYSADRSRITLDDPLSALDAHVNHIFNSAIRKHSKYVLF	739
Db	736	ERGVNISGGQORVSMARAVYSNSDVIYFDPPRLSALDAHVAHQVFDSCMKDELNGKTFVL	795
Qy	740	VTHOLOVYLVDCDEVIEKMEGCTTERGTHEEL-----NANLNDY	777
Db	796	VTNOLHFLPLMDKILIVESEGMKEEGTLEVELSGILFKKIMENAGKADATQVETNDE-	854
Qy	778	ATIFNNLLGEPPEINSSKKEETSSOCKKODKGKTISSIKKEKAYKREBEOQLQLEKG	837
Db	855	-----NIIKLGPIVTVDSER--NLGSTKOG-----KRRSV-----LIKQBERE	892
Qy	838	QGSVWSVSYGVYIQAGSGPLAFVYIMALFEMLVNGSTAFSTWMLSWYIKOGSGNTVTYRGN	897
Db	893	TGIIEMNVIAMRKKEAVGGLMYMILLACTIATLEVLRVSSWMLSTWTDQ-----	941
Qy	898	ETSVSDSKMDPEHMOYASIALSMAVMILKAIKRGVVEKGTLRASSRLHDELFRILLR	957
Db	942	-----STKSNVSPGFYIYVYVALLGFGQVAVYFTNSFWLITSSLHAARLHDAMLSILR	995
Qy	958	SPMKFDTPTPGRIILNRKSKMDDEVYVLRPQAEKFIQ---NVLLPFCVGMIGVPEWF	1014
Db	996	APMLEFHNPGRVYINRSKSDIGIDRRVANIMNNFNQMLQJLSTFALIGTVSTISLW-	1054
Qy	1015	LVAVPELVILVESVLIYERVLIRELRKLDNTQSPFLSHITSSIOGLATIHAYNGOEF	1074
Db	1055	---AIMPILILFPAALVLYQSTRSREYRRLDSYTRSTIYQOFGEALNGLSITAY-KAYIRM	1111
Qy	1075	HRYOELLDNOAPPELF--TCAMRWLAVRLDLISALITTTGMLVYMHGQIP---PAY	1128
Db	1112	AKINCKSDMNNIRFTLANTSSNRWLTIRLETGLGVMILTFATFAYLOGNNTNNOAGFAST	1171
Qy	1129	AGLAYSANQVLGLFQFVYRLASFEARFTSYERINNHITKLSLEAPARKNKAPSPMP	1188
Db	1172	MGLLSTYLTNTSLSGVLRQASRAENSLNSVERGNTI-DIPSEADITIEHNRVYCGMP	1230
Qy	1189	OGEVTFENAEMYREBNPLVLYKVSFTIKPREKIGIYGTGSGSGLSGMLFRLVELSG	1248
Db	1231	SGGSIKFFEDVHLARYRPGLPVYLHGILFEVPSSEKGVYGRGAGGSSMLNMLFRIVEEK	1290
Qy	1249	GCITKIDVARIISDGLADLRSKLSIIPQEPVLFSGTVRSNLDPEFNQYTDQIMDALERTHM	1308
Db	1291	GRIMIDDCDVAKFGLTDVRYLSIIPQSPVLFSGTVRENNIDPFSEHNDAGMEALHRAHI	1350
Qy	1309	KECIAQLOLKLKESVEMNGDNSVGEORLLCIARLLHCKIILIDEPTAMDEFTDILI	1368
Db	1351	KDVISRNPFGDAEVECEGENSVGQROLISLARLLRSRKILVYDEVTAAVADVRTDLSI	1410
Qy	1369	QETIREADACMLTIAIRLHTVLGSDRIIMVLAQOVVEFPPTVSLLSNDSSRFYAM---	1425
Db	1411	QRTIREEKSCHMLVIAIRLNTIIDCKILVYSSQVLEYEYDPSQBLSRDTSAPFRMWS	1470
Qy	1426	-----FAAENKVAVKG	1437
Db	1471	TGPANAQYLSNLVFERRENGMSVGG	1495

RESULT 8
D86428
hypothetical protein AAG28284.1 [Imported] - Arabidopsis thaliana
C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C.Accession: D86428
R.Thelogiadis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000
A.Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltil, R.; Marzla
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: D86428

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 11622 <STO>
 A:Cross-references: GB:AE005172; NID:g11055814; PIDN:AAG28284.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: human multidrug resistance protein cMOAT2; AMP-binding cassette homology

Query Match 27.7%; Score 2023.5; DB 2; length 1622;
 Best Local Similarity 34.4%; Pred. No. 1.5e-127;
 Matches 478; Conservative 239; Mismatches 500; Indels 171; Gaps 21;

QY 74 EENHKGTHHGLSLAKRTTCKHQHVDNAGLFSCMTFSWLSLAVAKKGLMEDV 133
 DB 217 EELPGGE-----NICEPRH---ANLEDSIFSWLNPMLTSGSR-PLTEKDV 259
 QY 134 WSLKHESSDVNCRRLRLMOELENVGPDAASLRVWVIFCPTRLLSIYCLMITOLAG 193
 DB 260 WHLDPTWKTETLWRSFQKSMDELEKRP-----W-----LIRALNSLIG 300
 QY 194 -----FSGPAFWYKHLETTQATES--NLQYSLLVGLLLEIVRSW 234
 DB 301 RFWMGFWKIGNDSCQFQV-LLNELLSQMLNEPAMIGYVAISIFVGVLCVE-- 357
 QY 235 SLALVTLANTRGVRLGAILITMAFKRIKLKNIKESL--GELINICSDNQMFEEAA 292
 DB 358 --AOYFQNMVWVGRILRSALIAVERKSLRTNIGRRKFPOTGKTNTMTDAESLQIQ 415
 QY 293 VGSILAGPVVALIIGTYNYVILGPTFLGSAVILFYPMAMFASRLTAVERRCVATD 352
 DB 416 SLHFMASPRRIIYALVLYQQLGVAIIGLFLVIMFPQIOTVLIISKTQKILRGLORTD 475
 QY 353 ERVOKNNEVLTYIKFIKVMYAVKAFSOSVOKIREERRILKAGYFOSITGVAPIVVI 412
 DB 476 KRIGLMEVLLAAMDPTKCYAMENSFOGKQVTRDELSEWFKQQLLSAFNFIINSIPVL 535
 QY 413 ASVYTFEVMHTLGGDLTAAQAFTVYVFNSTFALKATPPSVKSLASAVAVDFKSLFL 472
 DB 536 VTVAVFQVFLGLGDLTPARAFSLSLFVSILFPLFMIPIITQMVANAVSLNLEVLIS 595
 QY 473 MEVHMKNKPASP-HIKTEKKNATLAWDSHSSIQNSPKLPKMKDKRASRGKKEVR 531
 DB 596 TEERVLLPNPPIEGGPAISIRNCFESWD----- 625
 QY 532 QLORTEHQAVALBQKHLLSDERPSPREEEGKHILGHLRLQRTLSIDLETQEGKLV 591
 DB 626 -----KADRP-----TISNINLIDPLSLV 645
 QY 592 GIGGVSQGTSLISALIGOM-TLLEGSIAISGTFVAVAOAMLTNLTDRNLIIFGEYD 650
 DB 646 AAVGSTEGETSLISAMIGELPARSDATVTLRGSAVAVPOVSWTFNATVRNIIIFGAPFD 705
 QY 651 EERNNSVNLCCRLPDAIIPSSDLTEIGERGANLGGGQORISIALAVLSDRSITLDD 710
 DB 706 QEKTERYIDTALQHDLELLPGGDLTEIGERGVNISGGQKQKMSAAVAVSNSPVCILDD 765
 QY 711 PLSLDAHVGNHIFNSAIRKHSKYVLFYTHOLOYLVDDDEVIFEMKEGCTTERGTHEEL 770
 DB 766 PLSLDAHVGOQYFEKIKRELIGOTTRVLTNQLHFLSYQDKILLVHEGYKEGTYEEL 825
 QY 771 MNINGDIATIFNNL-----LIGETPVEINSKKETSSQOKSODKGPYTSIKKEKA 822
 DB 826 CHSGPLFQRLMENAGKVEDYSEENGAEVQTSYKPVENGANNMLQDKJETTKSGNS 885
 QY 823 VKRBEQGLVLEEGGSSVMSYGVYIOAAGPLAVLIVMALFMLVNGSAPSTWMLSY 882
 DB 886 Y-----LVKREKETGVSMKVLERYQNALGGAMVVMVLICVLTQVFRVSSSTWLSLSE 939

QY 883 WIKQSGNTVTYRGENETSVSDSKMDNPHMOYASITVLSMAVMLILKAIKGVFVKGLR 942
 DB 940 WTDGSG---TPKTHG-----PLEYNNIVYALLSPGVSVLINSYWLIMSLSY 982
 QY 943 ASSRLHDELFRILRSPMKPFEDTPTGRILNRFSKDMDEVDVRLPFOAEKIOVLIIVF 1002
 DB 983 AAKKMDMAGSLIRAPVVFQTPGLRIKRFKMDGDDIRYAVVNNMFSGIAQLLS 1042
 QY 1003 CVGMIAGVFWFLVAVGPLVILFVSLHIVSVRLIRELKRDLNITQSPFLSHITSSIGLA 1062
 DB 1043 TVILIGIVSTLSMAIMPLVAVFGAYLYQNTSREIKRMDOSTRSPVYAQFGALNGLS 1102
 QY 1063 TIHAYNKGOEFLHRYOELLDDNQAPFLFTCAMRWLAVRDLISALITTTGLMTIVMHG 1122
 DB 1103 SIRAYKAYDRMAEINSGRSMNNIFTLVNMAMNMLGIRLEVLGLMVLASLAVMQNG 1162
 QY 1123 QI--PRAVA---GLAISAVOGLTGFQTVYVLAETERTFVSERINHYIKTSLSEAPR 1177
 DB 1163 KANQAQASTWGLLSTSLTSLTSLTSLTSLTSLTSLTSLTSLTSLTSLTSLTSLTSLT 1221
 QY 1178 IKNRPSDPMFOEGEVTFENAEKRYRENTPLVKVSPFTIKPKKIGIVGTGSGKSSIG 1237
 DB 1222 IENRPPPGMPSGSSIKFEDVLYLRPELPVYLVHGVFSLSPMDKVGIVGTGAGKSSLL 1281
 QY 1238 MALRVLVLSGGCIKIDGVRIISDGLADLNSKLSIIPQEPVLSGTVRSNIDPPNYTDD 1297
 DB 1282 NALFRIVLEKGRILIDECIDIGRGLMDLRKRVLGITIQAPVLFSGTVRFNIDPSEHND 1341
 QY 1298 QIWDALERTHMKCEIAOLPLKLESEVMENGDNFSGVEGECIRALLRHCKITILDEAT 1357
 DB 1342 DWESLERAHKIDTIRRPFLGLDAEVTBAGENFSGVQRLSLRALLRKSILVIDEAT 1401
 QY 1358 AAMDTELDLIQETIRAFADCTMLTIAHLRLTVLSDRLVLAQGVVEEDPESVLLSN 1417
 DB 1402 AAVDVRDVLILQKTRIBREKSCYMLTIAHRLNTIIDCKVLVLDGKVGQERSPENILSN 1461
 QY 1418 DSRFPYAM 1425
 DB 1462 GESSFSKVM 1469

RESULT 9
 T01369
 ABC transporter AtMRP2 [imported] - Arabidopsis thaliana
 N:Alternate names: multidrug-resistance protein homolog T29F13.13
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001
 C:Accession: T01369; D84759
 R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K
 submitted to the EMBL Data Library, May 1998
 A:Description: Arabidopsis thaliana chromosome II BAC T29F13 genomic sequence.
 A:Reference number: 214179
 A:Accession: T01369
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1623 <ROU>
 A:Cross-references: EMBL:AC003096; NID:g3132469; PID:g3132479
 A:Experimental source: cultivar Columbia
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon,
 euss, D.; Nierman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
 Nature 402, 761-768, 1999
 A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: D84759
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1623 <STO>
 A:Cross-references: GB:AE002093; NID:g3132479; PIDN:ANCL6268.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g34660; T29F13.13
 A:Map position: 2

A:Introns: 110/3; 170/2; 183/3; 236/1; 274/2; 301/2; 309/3; 328/3; 350/3; 375/3; 411/3;
 C:Superfamily: human multidrug resistance protein cmoa2; ATP-binding cassette homology
 F:1259-1452/Domain: ATP-binding cassette homology <ABC>

Query Match 27.3% Score 1992; DB 2; Length 1623;
 Best Local Similarity 34.7%; Pred. No. 2e-125;
 Matches 478; Conservative 231; Mismatches 487; Indels 182; Gaps 27;

QY 100 PVDNAGLFSCFTFSLLSLAAVAAHKKGLSMEDVWSLSKHESDYVNCRRLEIMQDELNE 159
 D 100 PVDNAGLFSCFTFSLLSLAAVAAHKKGLSMEDVWSLSKHESDYVNCRRLEIMQDELNE 159
 D 227 PEKANINIDKIFSSWMNLMPLTSGSKR-PLTEKDVYLDTWQTEFLFTSFQSMWKELOK 285
 QY 160 VGPDAASLRVWVIFCRRLIISYICMLITOLAG-----FSGPAF--M 200
 D 160 VGPDAASLRVWVIFCRRLIISYICMLITOLAG-----FSGPAF--M 200
 D 286 PGP-----W-----LRLALNLSLGRFWMGFGFKIGNDOSQFPGPRLINDL 326
 QY 201 VKHLLFYQATESNLQYSLLLVGLLLEIYRSWSLALTWALNRYTGYRLRGALITMAFK 260
 D 201 VKHLLFYQATESNLQYSLLLVGLLLEIYRSWSLALTWALNRYTGYRLRGALITMAFK 260
 D 337 LKSMQEDAPAMWGYI-YAFSIFGVVFGVLCF---AQYFQWMEVGYRLRSALIAAYFR 381
 QY 261 KIKKAKNKEKSL--GELINICSDGQRMFEAAVGSLLAGPVAAILGMIVNIIIGPT 318
 D 362 KSLRLTNGRRRFGTGKTNLMTDAESLQICGSLHTMWSAPFRILIALILLYOOLGVA 441
 QY 319 GFLGSAVFILFYPAMFASRLTAYFRKCAATDERVOKMNEVLYIKFIKMYAWKAFS 378
 D 442 SLIGALLLVMPFLPLOTVLIISKQKLTREGLORTDKRIGLMNFVLAAMDVKCYIANENSFO 501
 QY 379 QSVOKIREERRILEKAGYFOSITGVAPIVVIVASVYFVSVMHTLGFDLTAQAFTVVT 438
 D 502 SKVQVRDDELDFWRFKSQLGLALNMFILNISPVLTVISFVGFTLLGGDLTPARAFSTLS 561
 QY 439 VENSMTFPAKATPEPVSKLSSEASVAVDRFKSLFLMEVYHMTKNKASP-HIKTEKNATL 497
 D 562 LEAVLRFLFPLFMTNITITOVANANSLKRLIEFLATEERLLPNPPIREPEAISISIRNGVF 621
 QY 498 AMSDSHSSIQNSPKLTPMKKDKRASRGKKEKVRQLQREHQVALAEOKGHLIDSDER 557
 D 622 SMDS-----K-----KG-----DHP 630
 QY 558 SPEEERKHHLGLRLQRTLSIDLEIOEGLVIGSGVSGKTSLSAIIQMTLLEG 617
 D 631 -----TLSINLMDVPLGSLVAVSGTGKTSLSAIIQMTLLEG 617
 QY 618 SI-AISGTFAYAAQQAQMLNATLNTLNIIIFGKTYDEERNVSNLSCCLPDLAIIPLSSDIT 676
 D 672 AIVTLRGSAVAYPOVSWIETAVRDNIILFGSPFREKERAIDVTSLKHDLELLPGDIT 731
 QY 677 EIGERGAMLSGGGQROKISLARALYSRSLYLLDPLSLADAVNGHINFSALRKHLKSKT 736
 D 732 EIGERGAMLSGGGQROKISLARALYSRSLYLLDPLSLADAVNGHINFSALRKHLKSKT 736
 QY 737 VLEVTYHQLQVLDCEVIFPKEGCITERGTBHELANLNNGDYATIFNNLLG----- 787
 D 792 RVLVYNQHLFLSQVDRIYLVHGGYKEGTYELSS-NG---PLQRLMENNKGVEYSE 847
 QY 788 ETPPYEINSKKE-----TSGSQAK-SODKGPKTGSIKKERAKVPEEGQIVOLEKGG 839
 D 848 ENGGEAEADQTEAQPANNGTNGIOMDGSDDKSKKGGKSGSV-----LIKQEREREG 901
 QY 840 SVPMVGVYVIOAAGPLAFLVIMALEMLANGSTAFSTWMLSTWYKOGSGNTVYRGNT 899
 D 902 VVSMVRLRYDQALGAVVVMMLLCYVLTVEYRVSSTWSEWDAG--EPKSHG-- 955
 QY 900 SVSDSMKDNPMHQYASITAYLSMAVMLILKARIGVAVFKGTARASSRLHDELFRRLINSP 959
 D 956 -----PLFVNLIIYALLTSFGQVLYTLNSYMLNLSIYAKKLHNDMLHSLIARP 1004
 QY 960 MKFPPTTPGRLINRFSKDMDEVDVRLPQAEMLIQNVILVEFCVGMIAAGVFPWELVAVG 1019
 D 1005 MSFETNPGLRIINRFARDLGDIDRTVAVFVWMFGVQSOLLSTVLLIGIVSTLSIMAIM 1064

QY 1020 PLVILFVSLHIVSRVLRLELKRDLNITOSPFLSHITSSIOGIATTIHAAYNGOEFLLHROE 1079
 D 1065 PLVILFVSLHIVSRVLRLELKRDLNITOSPFLSHITSSIOGIATTIHAAYNGOEFLLHROE 1079
 QY 1080 LID-----DNOAPPELFTC-AMRWLAVRDLISALTITTTGLMTVLMHQI--PPAYA- 1129
 D 1118 MADINGRSDNNIRFPLVVMGAMRNIGIRLETLGGLMTLWLTASFVVMONGRAENQOAFAS 1177
 QY 1130 --GLAISAVOLTGLEFQYTVRLASFTFARFISVERINIKTSLSEAPRIKNAKPSDPW 1187
 D 1178 TWGLISTALNTSLTGLVRLASLAEISLNAVERNGNTE-TPPEAPVIEINRPPGEM 1236
 QY 1188 PGEVTEFENAMERREMLPLVKKVSEFTIKPERKIGIYRTSGSKSISGMALFRLVELS 1247
 D 1237 PSSGSIKFEDEVLYLRPQLPVLHGVSFFIHPYDKVGIYGRGAKSKSLNMLFFIYVE 1296
 QY 1248 GGCIIKIDGVRISDGLADLRSLKSLIIPQEPVLPFGSTVNSNDPENQYEDQIWDALERTH 1307
 D 1297 KRRILIDCDVGKFGMLDKVGLIIPQSPVLPFGSTVRENDLPFGEHNDADLMESELERAH 1356
 QY 1308 MKECIAOLPKLSEVMEVNGDNFSGVERQLCIRALLRHCKIILIDEPATAMDETDL 1367
 D 1357 LKDTIRNDPLGIDAEVSEAGENSFVGQRLSLSRALLRSKYLVLDEXTAAVDVRTDAL 1416
 QY 1368 IOETIRFADCTMLTIAHRLTFLVGSPRIMVLAQGVVEFDPVSLSNSDSRFPYAM 1425
 D 1417 IOKTIRREKSGTMLIARLMTIIDCKIYLVDSGRVGFESFENLINSBSSFSKM 1474

RESULT 10
 S1863
 cadmium resistance protein ycf1 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YD9302.11c; protein YDR135c
 C:Species: Saccharomyces cerevisiae
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
 C:Accession: S51863; A55352; S50233
 R:Oliver, K.; Harris, D.
 submitted to the EMBL Data Library, February 1995
 A:Reference number: S51853
 A:Accession: S51863
 A:Molecule type: DNA
 A:Residues: 1-1515 <OL1>
 A:Cross-references: EMBL:Z48179; NID:965657; PIDN:CAA8217.1; PID:965668; GSPDB:GNO
 R:Schzykpa, M.S.; Wemmle, J.A.; Moye-Rowley, W.S.; Thiele, D.J.
 J. Biol. Chem. 269, 22853-22857, 1994
 A:Title: A yeast metal resistance protein similar to human cystic fibrosis transmembr
 A:Reference number: A55352; M01D:94357936
 A:Accession: A55352
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-679, R', 681-1515 <SZC>
 A:Cross-references: GB:LJ5237; NID:9556464; PIDN:AAA50353.1; PID:9556465
 C:Genetics:
 A:Gene: SGD:YCF1; MIPS:YDR135C
 A:Cross-references: SGD:S0002542; MIPS:YDR135C
 A:Map position: 4R
 C:Function:
 A:Description: required for cadmium resistance
 C:Superfamily: human multidrug resistance protein cmoa2; ATP-binding cassette homolo
 C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein; yeast vacuole
 F:238-308/Domain: transmembrane #status predicted <TM1>
 F:345-366/Domain: transmembrane #status predicted <TM2>
 F:421-442/Domain: transmembrane #status predicted <TM3>
 F:446-467/Domain: transmembrane #status predicted <TM4>
 F:533-555/Domain: transmembrane #status predicted <TM5>
 F:558-580/Domain: transmembrane #status predicted <TM6>
 F:646-829/Domain: ATP-binding cassette homology <ABC1>
 F:663-670/Region: nucleotide-binding motif A (P-loop)
 F:951-972/Domain: transmembrane #status predicted <TM7>
 F:995-1016/Domain: transmembrane #status predicted <TM8>
 F:1068-1088/Domain: transmembrane #status predicted <TM9>
 F:1092-1113/Domain: transmembrane #status predicted <TM10>
 F:1179-1200/Domain: transmembrane #status predicted <TM11>

F:1208-1229/Domain: transmembrane #status predicted <TM12>
 F:1289-1463/Domain: ATP-binding cassette homology <ABC2>
 F:1306-1313/Region: nucleotide-binding motif A (P-loop)

Query Match 27.18; Score 1983.5; DB 1; Length 1515;
 Best Local Similarity 33.18; Pred. No. 6,7e-125;
 Matches 472; Conservative 268; Mismatches 492; Indels 195; Gaps 27;

```

QY 77 PKGHHGLSALKPIRTCKHOPVDNAGLFCMTFMSLSLARVAHKKGLSMEDWSL 136
DB 196 PHQHIFHTLRRK-----NPYDSANIFSRITFTSMGLMTGYEK YLVADLTKL 246
QY 137 SKHESSDVCNRRLERLMOEELNEVGPDASLRVWIFCPT--RLILSTYCLMITOLAG 193
DB 247 PRNSESSELSOKLEKMNENELKOKSNPSLS-----MAICRTFGSKMLLAEPKAIHDVLA 301
QY 194 FSGP--AFWKHLEET-----QATESMLOYSLLVLDLL-----TE 229
DB 302 FTOQOLLRIILKFTVDYNSERODDHSSLOGFENHPQKLPVRGFLLAFAFLVGTQTS 361
QY 230 IVRSWSLALFWALNYRTGVRLGAILTMAFKILKNIKE--KSLGELINCSGDCQM 287
DB 362 VLHGYELN-----FNTGMITKSALTALITYOKSLVLSNEMSGSLSTGDIYMLAVDYOKL 416
QY 288 PEAAVGSLLAGPVAAILGKITVVIILPTGFGSAVFILFYPAMAPASRLTAYFRKC 347
DB 417 QDLTQWNLNLSWGPFQIILICLYSLKILGNSMVGVIILVLMPLNSFLMRIOKKLOKQ 476
QY 348 VAATDEVOKMNEVLYIKFKIMYAWKAFSOSVQKIR-EEERILKAGFOSITGVA 406
DB 477 MKYKDETRVISELNNIKSKILYAMKPYREKLEEVARNKELKNLTKLCYMAVTSPQF 536
QY 407 PIYVVASVTVFSYHM-TLGEDTLTAQAFTVTVVNSMTFALKVPFSVLSSEAVAD 465
DB 537 NIYVELVSCCTFAFVYTERBALTTDLVPLALTLFNLLSPMLITPWLNSIFASVIG 596
QY 466 RFLSFLMEVEYHMKNNKPASPHIKTEKKNATLAMDSSHSSIQNSPKLTPKKKKRARSRG 525
DB 597 RLTFEENEL-----OPDSVQRLPKY-KNIGVAINIG 629
QY 526 KKEKRVOLTEHOAVLAEOKGHLLDSDERPSPEEKGHIHGLRLQTLSDIDEI 585
DB 630 DDAFTLMQKREPKY-----ALKNINPQA 653
QY 586 QEGKLVIGSGVSGSGKTSLSAILGQMTLEGSIAISGTFAYVAQAQMLNATLRDILF 645
DB 654 KKGMLTCTIVGKVGSGTALLSCMLGDLFRVKGFATVHGSVAYVQVPMINGYKEMILF 713
QY 646 GKEVDERYSVLNSCCLRDLALTPSSDLTEIGERGANISGGQRQRISTARALYSRSTI 705
DB 714 GHRDADFERYEKTAKCALITDLALMDGDKTLVGEKGISLSGGKARISLARAYARADT 773
QY 706 YILDDPLSALDAHNGNIHFNSAIRKH--LKSQVLYFTVHOLOYVDDDEVIFMKEGCTE 763
DB 774 YLDDPLAADAENHARLIEHVLGPNLHRTKYVLAATNKVSLSTADSLTLNGELITQ 833
QY 764 RGTHEEL-NMLNGDYATIFENLLIGETPVEINSKETSQSGQKSDQKPKPTSGIKKEKA 822
DB 834 QGYVDETITKADSPMLKLNNY-----GKKNNG--KSNFEGDSSESVRESS 878
QY 823 VKPEEGLOQL-----EKGQGSVPWS 844
DB 879 I-PVEGLELDLOKINDLDFGNSDAISLRASDATLGSIDGDDENIAKRREREGQKAWN 937
QY 845 VGYVYIOAGGLAFIVYIMALFMLNGSTAFSTWMLSYWIKQSGNGTTVYRGNTSVDSDS 904
DB 938 IYLEYAKAC-NPKSVCYFIFLYISMLSVGNWMLKH-----SEVNSR 981
QY 905 MKDNPHQVYASIV--ALSMAVMLILKAIKGVVYVYKGLRASSRLHDELFRRLIRSPMK 961
DB 982 YGSNPNARALATAYFALGIGSALATLTIOTI-VLWVCTIHASKYLTNMLTNSVLRAPMT 1039

```

```

QY 962 FETDTPGRLINRFSKMDDEVDRLPFOAEMFIONVILVEFCVGMAGVPEWLVAVGPL 1021
DB 1040 FFETTPIGRLINRFSNNDIYKDALIGRTFSQFVNAVAKVFTPIVATWQFIITLPL 1099
QY 1022 VILFSLHIVSRVILRELKRLDNTOSPFSLSHITSIOGLATIHANYKGOEFLRYOGL 1081
DB 1100 SVFTIYVQOYLRSTRRELRLDSTRSPISYSHFQETLGGTLVGVGSOQKRFSHINCRI 1159
QY 1082 DDNOAPFELFCAMRWLAVALRDLISALITTTGLMIY-LMHGOIPRAYGLAISYVOL 1139
DB 1160 DNMSAFPSINARWLAVALRELIGSITILGAATLSVFLRKQGLFTGMVGLSLSYALQI 1219
QY 1140 TGLFOFTVRLASEPFAFTSYERLNHYIKTILSLAPARIKKAASPMQPOGEVTFENAE 1199
DB 1220 TQTLNWTIVRMTEVEETNIVSERIKER-ADLKSSEAPLIVEGHRPPKMPSPQDIFKNYS 1278
QY 1200 MRYRENPLVKKVSFTLKPREKIGTIGRGSGSSISGMALFRLVELSGGCIKIDGVRIS 1259
DB 1279 TRYPELDLVKHNINHKPREKVGIVGRGAGKSSLTALFRLIEASEGIVIDNAIN 1338
QY 1260 DIGLADLRSKLSIIPQEVLFSGTVRSNLDPEFNQYTEDQIDALERTHMKCIAQLPLK 1318
DB 1339 EIGLYDRHKLSIIPQDSQVREGTVRENIDPINOYTDEAIRALELSHKELHVLMSNDG 1398
QY 1319 LSEVMENGNFVSGERQLCTIARALRHCKILLDLDEATAAMDTEILLIOETIRAFAD 1378
DB 1399 LDAOITREGGMLVSGORQLCLARALVPSKTIIVLDEATAVDEETKRVQDETIRTAEKD 1458
QY 1379 CTMLTIARLHTFVGSRIWVLAQOVVEFPTPSVLSNDSRFPAM 1425
DB 1459 RTIITLIRLNTIMDSRIIVLDNGKVAEPDSQGLSDNKSLEYSL 1505

```

RESULT 11
 JE0336
 canalicular multispecific organic anion transporter - human
 C:Species: Homo sapiens (man)
 C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 04-Jun-1999
 C:Accession: JE0336
 R:Uchimi, T.; Hinoshita, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Toh, S.; Furukawa, Biochem. Biophys. Res. Commun. 252, 103-110, 1998
 A:Title: Isolation of a novel human canalicular multispecific organic anion transporter.
 A:Reference number: JE0336; MUID:99032812
 A:Accession: JE0336
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1527 <CDS>
 A:Cross-references: GB:AF083552
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homolo
 C:Keywords: ATP
 F:1306-1499/Domain: ATP-binding cassette homology <ABC2>

Query Match 26.64; Score 1941; DB 2; Length 1527;
 Best Local Similarity 32.28; Pred. No. 4.9e-122;
 Matches 463; Conservative 268; Mismatches 511; Indels 196; Gaps 22;

```

QY 85 LSAI-----KPIRTCK-----HOPVDNAGLFCMTFMSLSLARVAHKKGLSMEDV 133
DB 182 LSAIILACFRKRPFFSAKNVDPNPYPETSAGFLSRIFFWFTKMAIYGRH-PLERKDL 240
QY 134 WLSKHESSDVCNRRLERLMOEELNEVGPDAS-----L 167
DB 241 WLSKEEDRSQVYQQLLEAWRKQKQTAARHAKSAAPKNAAGSEDEVLLGARPRRPKPSFL 300
QY 168 RRVWICRRRLIISTYICMLITQLAGSGAPFVWKHILEYTOATESMLOYSLLVLDLL 227
DB 301 KALATP-GSSEFLISACFKLIQDLISFTNPQ-LTSLIRFISNPMGSGWGFVAGLMFL 358
QY 228 TEIYRWSLALFWALNYRTGVRLGAILTMAFKILKLN-IKESK-LGELINCSNDGQ 285
DB 359 CSMMQSLILOHYHYITVTVGKFRGTGIMGVYIRKALVITNSVKRASTVGEIVLMSVDAQ 418

```


QY 773 -----LNGDYATIFNNLLGSETPPVE-----IN 795
Db 890 RRRREAEASADDEDDENSEPFGIMIGSDSDEYDDDDVMA-SPLITDVLGTSHTSVSGIIN 948
QY 796 SKKETSGOKSQDQPKTKSIKKEKAVPEEGOLVLEKEKOGSVPMVGYVYIOAGG 855
Db 949 RRRISTSTHKORRLSTSTSHSTSTASSTORQLGTRVEFGKMKPTYYKFFCAMOM 1008
QY 856 PLATFVLMALFMLNAGSTAFS---TWMISYWKQSGNTTTRGNETSVSDSKMDPNHQ 912
Db 1009 STAVLEFLVGM-----TSTSTFSGMRNMLTFDW-----SNDNAARSGSNTT-----GQPIA 1053
QY 913 YVASIYALSMAMMLIKARGVFVKGTIRASSRLHDELFRILRSPMKFPDTPTRGLL 972
Db 1054 IRLGAYAGIGFSEIILLTGLMISLLYGGVASRNLHAPMLRMLFRYPMAFYDTPTRGLL 1113
QY 973 NRRSKDMDEVDRLPFOAMFIQNVILVFCVGMAGVPMFLVAVGPVILFESVLHYS 1032
Db 1114 NRIKGDIEIVDVLPPNVOFGACLOVSTLIIMISTPVGIVYIPLSVATLMRMRY 1173
QY 1033 RVLBELKRLDITOSPFLSHSTSIQIATTHAYNKGOEFLHRYQELLDDNOAPEFLFT 1092
Db 1174 IATSRQLKRESITRSPISYSHLSSESIOGSATIRAYHLVDRFCKLSETKYDSHVQCRILNY 1233
QY 1093 CAMRNLAVRLDIALITLITGLMIVLMHGOIPRAVAGLISAVOGLGLOFTVPLASE 1152
Db 1234 VANRMLSVLSEIFGNCITVLSALFPALPRTTTTSGVIGSVSALITVTLNPAVQITRK 1293
QY 1153 TEARFTVERINHVIKTLSLEAPARIKN---KAPSPDMQEGEVTPENAEEMRYENLPV 1209
Db 1294 LETNIVSVRVEAEYAEET---ETEAEMKSEPKGEPQNMVSEGRIVANNNSARIRPLNIV 1350
QY 1210 LKKVSETTKPKKIGIVGTGSGKSSLSGALFLVELSGGCIKIDGVRISDGLADLRSK 1269
Db 1351 VQOLNVEIKPHEKVGIVGTGAGKSSVTLSPRIIEAAGQIVYDGINIAEIGLHDLRSN 1410
QY 1270 LSIITPOEPLFGTVASNLDPFNOYTEDQIMDALERTHKECTAQPLPLESEVMEGND 1329
Db 1411 LTTITPDVPLESGTTLRFNLPDPPNNHSDGIMKTLBMANKEPRTAHNEOLNITTEGGN 1470
QY 1330 FSVGEROLLICARALLRHCKILILDEATAMPDETDLTIQETIREAFADCTMLTIAHRL 1389
Db 1471 ISVGQRQLVCARALLRKRRVILILDEATAVYSDTALQIKTIRREFANATVLTIAHRLN 1530
QY 1390 TVLGSDRIMVLAOGOVEEDTPSVLLSNDSSRPYAM 1425
Db 1531 TTMDDYRITIVLNDGKVGEDSPANLISNRNSEYSM 1566
RESULT 14
T48059
ABC transporter-like protein - Arabidopsis thaliana
N:Alternate names: protein F26K9.130
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 21-Jul-2000
R:Accession: T48059
R:Biocheck, H.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.; Quettier, F.; Salanoubat
submitted to the Protein Sequence Database, March 2000
A:Reference number: 224465
A:Accession: T48059
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1539 <BL0>
A:Cross-References: EMBL: AL162651
A:Experimental source: cultivar Columbia; BAC clone F26K9
C:Genetics:
A:Map position: 3
A:Introns: 804/3; 833/3; 1078/3; 1133/3; 1232/1; 1405/3; 1427/1; 1507/1
A:Note: F26K9.130
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 25.8%; Score 1882; DB 2; Length 1539;
Best Local Similarity 32.2%; Pred. No. 4.7e-118;
Matches 454; Conservative 277; Mismatches 485; Indels 194; Gaps 30;
QY 102 DNAGLFSCTMF-----SWLSLARVAHKKGELSMEDVSLSKHSSDYNCRLERLW-Q 154
Db 245 ENVSILASASFSISKTPWLMNPNLRRKGY-KSPLNLDQVPTLSEHRAEKATLTFESKMPK 303
QY 155 EELNEVGPDAASLRVRYWTFCTRLILSTVCLMTIQLAGSPAPFVKKLLLETQATBSN 214
Db 304 PONSNNPVRTLLIRCFMWEIAFTVALAIRLSVI-----YVGPV-LIOSFVFTSGKRS 358
QY 215 LQYSLLVYGLLTTELVRMSLALFWALNY---RQVRLGALITWAPFKILIKNIKER 271
Db 359 PEOGYTLVILILAKFEVLS---THQFNNSOKGLMLIRSLITLALYKKGKLITGSARO 415
QY 272 S--LGLINICNSDQGR---MFEAAVGSLLAGFPVAIILGIVNVIILGPT-----G 319
Db 416 NHGQGVIVYMAVDAGQLSDMLQLHAI--WMLPLQVAALVLLYNT--LGPVTVTVIG 471
QY 320 FLGSAVFLIFYPAMFASRLTAFRRKCYAADDERVQKANEVLTLYKFKIMAYAKFSG 379
Db 472 LGIGIFVFTL-----LGTRRNNRYOPSLMNNRDSRRKATNEMLVNRYIKFOAMEDHNE 525
QY 380 SVQKIREERRLIEKAGYQOSITGVAPIVVYIASVYFVSVHMTLGEFDLTAQAFTVTVY 439
Db 526 RILKREMEFGMLSKFLYSIAGNIYLMSTPVLSALITTTAVFLQVADAGVFTTTTI 585
QY 440 FNSMTPALKVTPSVKSLSEASVADRFKSLFMEVH---MINKNPASPHIKIEKNAT 496
Db 586 FKIQGPIRTPQSMISLSQAMISGLRDAYMMSRLESEYERSGCGGONAVEIKDS 645
QY 497 LAMDSHSSITQNSPKLTPMKKDKRASRGKKEVROLQTRHOAVLAEOKGHLLDSER 556
Db 646 FNSD-----DEDE 654
QY 557 PSPREEEGKHILHRLQRTLHSDILEIOEGLVIGCSVSGSKTSLAISILIQMTLE 616
Db 655 PA-----IENINFEVAKGELAVIGTVSGKSSLASVAGEHKLIS 695
QY 617 GSIAISGTFAYAAQAMILNATLRLDNLIFGEKEYDEERYNSVNSCLRPDLALPSSDLT 676
Db 696 GKRVCGTAAVYAQTSWIONGVQDNLILFGLPMNRSKYNEVLKVCLEKDMQMEGGDT 755
QY 677 EIGERANISGGQRORISLARALYSRSTIYIDPLSALDAHGNHIFSAIRKHLKST 736
Db 756 EIGERINISGGQRORISLARALYSRSTIYIDPLSALDAHGNHIFSAIRKHLKST 815
QY 737 VLEFVTHQLYVDCDEVIFMKEGCTERGGHEELMLNGDYATIFNNLLGSETPP--VET 794
Db 816 ILVTHQVDFLHNVDKILVNRDMQIVQSGKIDELVSSGLDFGL--VAAHETSMELVER 872
QY 795 NSKKESTGS-----QKKSODKGRK--TGSIKKEKAVK----- 824
Db 873 GSASATAANVPMASPTQRSISTIESPROKSPKVRHRTSMESPRVLRRTSMESPRLELN 932
QY 825 -----PEEG-QVQVLEKEKOGSVPMVGYVYIOAGGRLATFVLMALFMLNAG 871
Db 933 DESIKSFILSNIPEDSGRLTIKEEREVGQVSFOVYLYLSTEALGWMGMLILVFFSVAMQA 992
QY 872 SFAFSYMWLSYWKQSGNTTTRGNETSVSDSKMDPNHQYVASIYALSMAMMLIKAI 931
Db 993 SLMASDYWLAY-----ETSAKNVSDAYV--FIRYVIAAIVLYLCL 1035
QY 932 RGVVFAVKGTLRASSRLHDELFRILRSPMKFPDTPTRGLIRFSDKMDDEVDRLPQ-- 989
Db 1036 RAYVYTHLGKTAQIFPKQILNLSLVHAPMSFPTPSGRILSRASDQNVNVDIFIPMIG 1095
QY 990 --AEMFIQNVILVFCVGMAGVPMFLVAVGPVILFESVLHYSVNLIREKRLDITNQ 1047
Db 1096 LVATNM--TLLSLTIVTCQYAMPVTFEILPLGLWLNWYKRYVLAAS--RELRLDITSTK 1151
QY 1048 SPLSHITSSIOGLATTHAYNKGOEFLHRYQELLDDNOAPFLFTCAMRNLAVRLDLIS 1107


```
Db 1281 TLTFLALFRLEPTSGDIQUDPINITSIGIHLDRSLATIPQENQAFEGTIRENLDPNANA 1340
QY 1295 TEDQIMDALERTHMECTAQLPLKLESEYMENGDNFSYGEROLLCIARALLRHCKILIID 1354
Db 1341 TDEEIWHALEAASIKQFOTLDGGLYSRYTEGAMISSGOROLMCLTRALITPTREVLLID 1400
QY 1355 EATAMDETDLIOETIREAFADCTMLTIAHRLHTVLGSDRIWVLAOGVVEFDTPSVL 1414
Db 1401 EATPAVDVEIDAIYDRTIREFNNDKITLITAHKINIVMDSNRILVLDHGKVVVEFDSTKKL 1460
QY 1415 LSNDSSREFYAM 1425
Db 1461 LENKASLFYSL 1471
```

Search completed: November 9, 2001, 17:48:05
Job time: 6553 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 17:48:02 ; Search time 78.74 Seconds
(without alignments)
625.160 Million cell updates/sec

Title: US-09-528-031-2
Perfect score: 7308
Sequence: 1 MKDIDIGKEYLIPSPGYRSV.....DSRFYAMFAAEKNKVAVKG 1437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7297	99.8	1437	1	MRP5_HUMAN
2	6923.5	94.7	1436	1	MRP5_RAT
3	6855.5	93.8	1436	1	MRP5_MOUSE
4	2100.5	28.7	1541	1	MRP2_RAT
5	2082	28.5	1531	1	MRP1_HUMAN
6	2073	28.4	1545	1	MRP2_HUMAN
7	2053.5	28.1	1325	1	MRP4_HUMAN
8	2033.5	27.8	1564	1	MRP2_RABIT
9	1983.5	27.1	1515	1	YCF1_YEAST
10	1960	26.8	1522	1	MRP3_RAT
11	1945	26.6	1527	1	MRP3_HUMAN
12	1874.5	25.6	1478	1	YAMB_SCHPO
13	1759	24.1	1580	1	ACCB_HUMAN
14	1749	23.9	1592	1	YHDS_YEAST
15	1742.5	23.8	1581	1	ACCB_CRICR
16	1733.5	23.7	1559	1	BPT1_YEAST
17	1732	23.7	1580	1	ACCB_RAT
18	1727	23.6	1477	1	YOR1_YEAST
19	1700.5	23.3	1503	1	MRP6_HUMAN
20	1665	22.8	1502	1	MRP6_RAT
21	1649.5	22.6	1661	1	YPR1_YEAST
22	1592	20.8	1548	1	MDR1_YEAST
23	1451	19.9	1492	1	CFTR_SQUAC
24	1408	19.3	1480	1	CFTR_HUMAN
25	1405.5	19.2	1558	1	YK83_YEAST
26	1381	18.9	1481	1	CFTR_SHEEP
27	1378	18.9	1476	1	CFTR_MOUSE
28	1369	18.7	1481	1	CFTR_BOVIN
29	1365.5	18.7	1485	1	CFTR_XENTIA
30	1334.9	18.5	1427	1	ABCL1_SCHPO
31	1329.5	18.2	1450	1	CFTR_RABIT
32	785.5	10.7	1276	1	MDR2_CRIGR
33	774.5	10.6	1321	1	MDR1_CAEEL

34	767.5	10.5	1280	1	MDR1_HUMAN	P08183 homo sapien
35	766.5	10.5	1276	1	MDR1_CRIGR	P21448 cricetus
36	764.5	10.5	1276	1	MDR1_MOUSE	P06795 mus musculus
37	764.5	10.5	1281	1	MDR3_CRIGR	P23174 cricetus
38	759.5	10.4	1278	1	MDR2_RAT	P08201 rattus norv
39	754.5	10.3	1277	1	MDR1_RAT	P43245 rattus norv
40	752	10.3	1276	1	MDR2_MOUSE	P21440 mus musculus
41	750.5	10.3	1279	1	MDR3_HUMAN	P21439 homo sapien
42	740.5	10.1	1302	1	MDR4_DROME	P00449 drosophila
43	735.5	10.1	1276	1	MDR3_MOUSE	P21447 mus musculus
44	656.5	9.0	1362	1	PNL1_SCHPO	P36619 schizosacch
45	638	8.7	1290	1	STE6_YEAST	P12866 saccharomyc

ALIGNMENTS

RESULT 1
ID MRP5_HUMAN STANDARD; PRT; 1437 AA.
AC 015440: Q9UOC3; Q9UNP5; Q9UN85; 014517;
DT 15-JUL-1998 (Rel. 36, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MULTIDRUG RESISTANCE-ASSOCIATED PROTEIN 5 (ABC TRANSPORTER MOAT-C)
DE (PABCL1) (SMRP).
GN ABC5 OR MRP5.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99043202; PubMed=9827529;
RA Belinsky M.G., Bain L.J., Balsara B.B., Testa J.R., Krueh G.D.;
RT "Characterization of MOAT-C and MOAT-D, new members of the MRP/CMOAT
RT subfamily of transporter proteins.";
RL J. Natl. Cancer Inst. 90:1735-1741(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20184734; PubMed=10721709;
RA Suzuki T., Sasaki H., Kuh H.J., Agui M., Tatsumi Y., Tanabe S.,
RT Terada M., Saijo N., Nishio K.;
RT "Detailed structural analysis on both human MRP5 and mouse mrp5
RN proteins, has anion transporter activity but does not confer multidrug
RN resistance when overexpressed in human embryonic kidney 293 cells.";
RN J. Biol. Chem. 274:23541-23548(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20184734; PubMed=10721709;
RA Suzuki T., Sasaki H., Kuh H.J., Agui M., Tatsumi Y., Tanabe S.,
RT Terada M., Saijo N., Nishio K.;
RT "Detailed structural analysis on both human MRP5 and mouse mrp5
RN proteins, has anion transporter activity but does not confer multidrug
RN resistance when overexpressed in human embryonic kidney 293 cells.";
RN J. Biol. Chem. 274:23541-23548(1999).
RN [4]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=20319051; PubMed=10840050;
RA Wijnholds J., Mol C.A.A.M., van Deemter L., de Haas M., Scheffer G.L.,
RT Baas F., Beijnen J.H., Schepers R.J., Hulse S., De Clercq E.,
RN Balzarini J., Borst P.;
RT "Multidrug-resistance protein 5 is a multispecific organic anion
RT transporter able to transport nucleotide analogs.";
RN Proc. Natl. Acad. Sci. U.S.A. 97:7476-7481(2000).
RN [5]
RP SEQUENCE OF 492-1437 FROM N.A.
RX MEDLINE=97472289; PubMed=9325169;
RA Suzuki T., Nishio K., Sasaki H., Kurokawa H., Saito-Ohara F.,
RT Ikeuchi T., Tanabe S., Terada M., Saijo N.;
RT "CDN cloning of a short type of multidrug resistance protein
RT homologue, SMRP, from a human lung cancer cell line.";

BL Biochem. Biophys. Res. Commun. 238:790-794(1997).
[6]
RN SEQUENCE OF 1216-1437 FROM N.A.
RP TISSUE-Brain;
RC MEDLINE-97413640; PubMed-9270026;
RA Kool M., de Haas M., Scheffer G.L., Scheper R.J., van Eijk M.J.,
RT Juljn J.A., Baas F., Borst P.;
RT "Analysis of expression of cMOAT (MRP2), MRP3, MRP4, and MRP5,
RT homologues of the multidrug resistance-associated protein gene
RT (MRP1), in human cancer cell lines";
RL Cancer Res. 57:3537-3547(1997).
CC -1- FUNCTION: ACTS AS A MULTISPECIFIC ORGANIC ANION PUMP WHICH CAN
CC TRANSPORT NUCLEOTIDE ANALOGS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MRP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF104942; AAD04169.1; -
DR EMBL; AF1046074; AAD37716.1; -
DR EMBL; AB019002; BAA76608.1; -
DR EMBL; AB005659; BAA22887.1; -
DR EMBL; U83661; AAB71758.2; -
DR InterPro; IPR001140; -
DR InterPro; IPR001617; -
DR Pfam; PF00664; ABC_membrane; 2.
DR ProSite; PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat.
KW TRANSMEM 179 POTENTIAL.
FT TRANSMEM 219 239 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 400 420 POTENTIAL.
FT TRANSMEM 434 454 POTENTIAL.
FT TRANSMEM 608 628 POTENTIAL.
FT TRANSMEM 848 868 POTENTIAL.
FT TRANSMEM 917 937 POTENTIAL.
FT TRANSMEM 997 1017 POTENTIAL.
FT TRANSMEM 1018 1038 POTENTIAL.
FT TRANSMEM 1104 1124 POTENTIAL.
FT TRANSMEM 1127 1147 POTENTIAL.
FT NP_BIND 595 602 ATP (POTENTIAL).
FT NP_BIND 1227 1234 ATP (POTENTIAL).
FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 684 684 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 890 890 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 897 897 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1044 1044 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1329 1329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1417 1417 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 176 176 R -> P (IN REF. 3).
FT CONFLICT 400 400 S -> G (IN REF. 1).
FT CONFLICT 581 581 I -> V (IN REF. 2).
FT CONFLICT 1383 1383 T -> N (IN REF. 3 AND 5).
SQ SEQUENCE 1437 AA; 160659 MW; 00558076B3BB4C00 CRC64;
Query Match 99.8%; Score 7297; DB 1; Length 1437;
Best Local Similarity 99.98%; Freq. No. 0;
Matches 1435; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 61 LDASHNSQLRLIDEEHPKXKHHGSLALKPIRTTCKHOPVDNAGLSQMTFSWLSLAR 120
DB 61 IDASHMSQLRLIDEEHPKXKHHGSLALKPIRTTCKHOPVDNAGLSQMTFSWLSLAR 120
QY 121 VAHKKGELSMEDVWSLSHSSSDVNCRLERLMOEELNEVGPDAASLRVYVWICRRLI 180
DB 121 VAHKKGELSMEDVWSLSHSSSDVNCRLERLMOEELNEVGPDAASLRVYVWICRRLI 180
QY 181 LSIYCLMTQLAGFSGPAPFVWKHLLEYQATESNLQYSLLVGLLLEIYRSMSLATW 240
DB 181 LSIYCLMTQLAGFSGPAPFVWKHLLEYQATESNLQYSLLVGLLLEIYRSMSLATW 240
QY 241 ALNRYTVRLRGAILTPAFAKKILKIKIKESLSELINICNSDGRMEFAAAGSLAGC 300
DB 241 ALNRYTVRLRGAILTPAFAKKILKIKIKESLSELINICNSDGRMEFAAAGSLAGC 300
QY 301 PVVALIGMIVNVLILGPGLGSAVFILFYAPMFAASRLTFRRKCYAANDEROKNE 360
DB 301 PVVALIGMIVNVLILGPGLGSAVFILFYAPMFAASRLTFRRKCYAANDEROKNE 360
QY 361 VLTYYKFKMTAWKAFQSQYOKIREERRLIEKAGYFQSIYGVAPLVVYIASVYFVS 420
DB 361 VLTYYKFKMTAWKAFQSQYOKIREERRLIEKAGYFQSIYGVAPLVVYIASVYFVS 420
QY 421 HMTLGFDLTAQAFVTVVFNSTPALKVTPSVKSLSEASAYANDRFSKSLMEVHIK 480
DB 421 HMTLGFDLTAQAFVTVVFNSTPALKVTPSVKSLSEASAYANDRFSKSLMEVHIK 480
QY 481 NKPAAPHIKIEKKNTLAMDSSHSSIONSPLTPKMKDKRASRCKKRYKQLOTEHQ 540
DB 481 NKPAAPHIKIEKKNTLAMDSSHSSIONSPLTPKMKDKRASRCKKRYKQLOTEHQ 540
QY 541 VLAEOGHLILDSDEPSPPEEKGHIHLGHLRLQRTLSIDLIEQEKLVGICGSVSG 600
DB 541 VLAEOGHLILDSDEPSPPEEKGHIHLGHLRLQRTLSIDLIEQEKLVGICGSVSG 600
QY 601 KTSLSAILGQMTLEGGSAISGTFAYVAQQAAMILNATLRDMLFGKRYDEERYSVNS 660
DB 601 KTSLSAILGQMTLEGGSAISGTFAYVAQQAAMILNATLRDMLFGKRYDEERYSVNS 660
QY 661 CCLRPDLAIPSSDLTEIGERGANLSCGORISLARALSPRSYIIDDPLSALDAHV 720
DB 661 CCLRPDLAIPSSDLTEIGERGANLSCGORISLARALSPRSYIIDDPLSALDAHV 720
QY 721 NHIFNSAIRKHLKSTVLFVTHQLQYLVDCEVIEFMKECCITERGTHELMNLNDYATI 780
DB 721 NHIFNSAIRKHLKSTVLFVTHQLQYLVDCEVIEFMKECCITERGTHELMNLNDYATI 780
QY 781 FNNLLGEPPEPEINISKETSGSKKQDKPKTGSIKKXAVKPEEQOLVLEKGGGS 840
DB 781 FNNLLGEPPEPEINISKETSGSKKQDKPKTGSIKKXAVKPEEQOLVLEKGGGS 840
QY 841 VPMASYGVYIOAAGPLAFVLMALFMLNVGSAFSTWMLSWIKOGSGNTVTGNETS 900
DB 841 VPMASYGVYIOAAGPLAFVLMALFMLNVGSAFSTWMLSWIKOGSGNTVTGNETS 900
QY 901 VDSMKNDPDMQYASIALSMAVMLILKALRGVVFVNGTLRASSRLHDELFRILRSPM 960
DB 901 VDSMKNDPDMQYASIALSMAVMLILKALRGVVFVNGTLRASSRLHDELFRILRSPM 960
QY 961 KFFDTTPGRIILNRSKMDDEVDRLPQAEFIONVILVFCVGMINGVPMFLVAAGP 1020
DB 961 KFFDTTPGRIILNRSKMDDEVDRLPQAEFIONVILVFCVGMINGVPMFLVAAGP 1020
QY 1021 LVILPSVLAHSRVILRELKRDNTQSPFLSHSTSIQGLATIHAYNKGOFELRHYOEL 1080
DB 1021 LVILPSVLAHSRVILRELKRDNTQSPFLSHSTSIQGLATIHAYNKGOFELRHYOEL 1080
QY 1081 LDDNQAPEFLFCAMRWLAVRDLISALITTTGLMIVLMHGOIPPAYAGLAISYAVOLT 1140
DB 1081 LDDNQAPEFLFCAMRWLAVRDLISALITTTGLMIVLMHGOIPPAYAGLAISYAVOLT 1140

QY 1141 GLEFOTVRLASEETAEPTVERINHYIKTSLSEAPARIKKKAPSPMPQCEVTEFENAEK 1200
 DB 1141 GLEFOTVRLASEETAEPTVERINHYIKTSLSEAPARIKKKAPSPMPQCEVTEFENAEK 1200
 QY 1201 RRENPLVLKRVSTIKRKERIGIVRGSGKSGMALFRLVELSGCICKIDGVRISD 1260
 DB 1201 RRENPLVLKRVSTIKRKERIGIVRGSGKSGMALFRLVELSGCICKIDGVRISD 1260
 QY 1261 IGLADRSKLSITIPQBPVLFSGTVRSNLDPEFNOYTEDQIWDALERTHMECLAQPLKLE 1320
 DB 1261 IGLADRSKLSITIPQBPVLFSGTVRSNLDPEFNOYTEDQIWDALERTHMECLAQPLKLE 1320
 QY 1321 SEVMENGDNFSGVEROLLIARALLRCKILLIDEATAAMDTELDLIQETIREAFADCT 1380
 DB 1321 SEVMENGDNFSGVEROLLIARALLRCKILLIDEATAAMDTELDLIQETIREAFADCT 1380
 QY 1381 MULTIAHRLFTVLGSDRIMVLAGOVVEPTPPSVYLSNDSSRFYAFMAAENKVAVK 1437
 DB 1381 MULTIAHRLFTVLGSDRIMVLAGOVVEPTPPSVYLSNDSSRFYAFMAAENKVAVK 1437
 RESULT 2
 MRPS_RAT STANDARD; PRT; 1436 AA.
 AC 090YM0;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MULTIDRUG RESISTANCE-ASSOCIATED PROTEIN 5.
 GN MRPS.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCBL_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
 RA Homma M., Suzuki H., Sugiyama Y.;
 RL Submitted (Nov-1998) to the EMBL/GenBank/DDJ databases.
 CC -1- FUNCTION: ACTS AS A MULTISPECIFIC ORGANIC ANION PUMP WHICH CAN
 TRANSPORT NUCLEOTIDE ANALOGS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 (ABC TRANSPORTERS). MRP SUBFAMILY
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AB020209; BAA88897.1; -;
 DR InterPro: IPR001140; -;
 DR Pfam: PF00064; ABC_tran; 2.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat.
 FT TRANSMEM 179 199 POTENTIAL.
 FT TRANSMEM 219 239 POTENTIAL.
 FT TRANSMEM 236 316 POTENTIAL.
 FT TRANSMEM 317 337 POTENTIAL.
 FT TRANSMEM 400 420 POTENTIAL.
 FT TRANSMEM 426 446 POTENTIAL.
 FT TRANSMEM 608 628 POTENTIAL.
 FT TRANSMEM 847 867 POTENTIAL.
 FT TRANSMEM 916 936 POTENTIAL.
 FT TRANSMEM 996 1016 POTENTIAL.
 FT TRANSMEM 1017 1037 POTENTIAL.
 FT TRANSMEM 1101 1121 POTENTIAL.
 FT TRANSMEM 1126 1146 POTENTIAL.

FT NP_BIND 595 602 ATP (POTENTIAL).
 FT NP_BIND 1226 1233 ATP (POTENTIAL).
 FT CARBOHYD 494 494 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 636 636 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 684 684 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 889 889 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 896 896 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1043 1043 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1328 1328 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1416 1416 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 1436 AA; 16085 MW; 10FE53B800531598 CRC64;
 Query Match 94.7%; Score 6923.5; DB 1; Length 1436;
 Best Local Similarity 94.7%; Pred. No. 0;
 Matches 1361; Conservative 34; Mismatches 41; Indels 1; Gaps 1;
 QY 1 MKDIDIGKEYIIPSPGYRSVREPTSTGTHRDREDSKFRRTRELECODALETAARAGLS 60
 DB 1 MKDIDIGKEYIIPSPGYRSVREPTSTGTHRDREDSKFRRTRELECODALETAARAGLS 60
 QY 61 LDASMSQRLIDEERPKKYHHGLSALKPRTTCKHQHPVDNAGLFSCFTESWLSLAR 120
 DB 61 LDIVSHSHLOIIDEERTKGYHHGLSALKPRTTCKHQHPVDNAGLFSCFTESWLSPLAQ 120
 QY 121 VAHKKGEISMEDVWSLSKHESSDVNCRRLERLMOEELNEVGPAAISLRVWVFCRTRL 180
 DB 121 VAHKKGEISMEDVWSLSKHESSDVNCRRLERLMOEELNEVGPAAISLRVWVFCRTRL 180
 QY 181 LSIYCLMTIQLAGSGPAFWVKHLEVTQATSESNLOYSLVLGLLTETIVRSMSIALTW 240
 DB 181 LSIYCLMTIQLAGSGPAFWVKHLEVTQATSESNLOYSLVLGLLTETIVRSMSIALTW 240
 QY 241 ALMYRTGVRLGAILTMAEFKIKLKNIKESYLGELINTCSNDGORMFEAAVAGSLIAG 300
 DB 241 ALMYRTGVRLGAILTMAEFKIKLKNIKESYLGELINTCSNDGORMFEAAVAGSLIAG 300
 QY 301 PVAALIGMIYNTIILPTGGLSVAFLIFYPAMMPASRLAYRRRCVAATDRVOKME 360
 DB 301 PVAALIGMIYNTIILPTGGLSVAFLIFYPAMMPASRLAYRRRCVAATDRVOKME 360
 QY 361 VLTIRFIKMYAVKAFSCQVQKIREERIRILEKAGFQSGITGVAPIVVIASVTFV 420
 DB 361 VLTIRFIKMYAVKAFSCQVQKIREERIRILEKAGFQSGITGVAPIVVIASVTFV 420
 QY 421 HMTLGFDLTAQAFTVYVFNSTFALKVTPESVKSLSASAVALDRFKSLFMEVHMIR 480
 DB 421 HMTLGFDLTAQAFTVYVFNSTFALKVTPESVKSLSASAVALDRFKSLFMEVHMIR 480
 QY 481 NKRPASHIKIEMKNATLAWSSHSSTONSFKLPKMKKKORASRGKKEKVRQIQRTEHOA 540
 DB 481 NKRPASHIKIEMKNATLAWSSHSSTONSFKLPKMKKKORASRGKKEKVRQIQRTEHOA 540
 QY 541 VLAEOKGHLILDSDEPSPPEEGKHILGHLRLQRTLHIDLEIOEGKLVGICSGVSG 600
 DB 541 VLAEOKGHLILDSDEPSPPEEGKHILGHLRLQRTLHIDLEIOEGKLVGICSGVSG 600
 QY 601 KTSLSIALIGQMTLLGSLAISGTFAYVAQAAMILNATLNDNLIFGKEYDEERYNSVLNS 660
 DB 601 KTSLSIALIGQMTLLGSLAISGTFAYVAQAAMILNATLNDNLIFGKEYDEERYNSVLNS 660
 QY 661 CCLRPDLAIIIPSSDLTEIGRGANLSGGQORISLARALYSDNSIYLLDPLSALDAHV 720
 DB 661 CCLRPDLAIIIPSSDLTEIGRGANLSGGQORISLARALYSDNSIYLLDPLSALDAHV 720
 QY 721 NHIFNSAIRKHLKSKTVLFTYHOLQYLVDCEVIFPKKEGCTIRGRGHEELMANGDPA 780
 DB 721 NHIFNSAIRKHLKSKTVLFTYHOLQYLVDCEVIFPKKEGCTIRGRGHEELMANGDPA 780
 QY 781 FNNLLGEPPEVINSKKESTSGSKSODKGPRTGSIKKKAKAVPEBQOLVQLEKGG 840
 DB 781 FNNLLGEPPEVINSKKESTSGSKSODKGPRTGSIKKKAKAVPEBQOLVQLEKGG 840
 QY 840 FNNLLGEPPEVINSKKESTSGSKSODKGPRTGSIKKKAKAVPEBQOLVQLEKGG 840
 DB 840 FNNLLGEPPEVINSKKESTSGSKSODKGPRTGSIKKKAKAVPEBQOLVQLEKGG 840

QY	841	VPMSPVGVYIOAAGGPAFLFYVIALFPLNNGSTAEFNMWLSYIKGSGMTVTYRGERS	900
Db	840	VPMSPVGVYIOAAGGPAFLFYVIALFPLNNGSTAEFNMWLSYIKGSGMTVTYRGERS	899
QY	901	VSDSMCKNPMHQAYIASYIALSMAVMILTKIRGVVFYKGLTRASSRLDELFRRLIRSPM	960
Db	900	VSDSMCKNPMHQAYIASYIALSMAVMILTKIRGVVFYKGLTRASSRLDELFRRLIRSPM	959
QY	961	KFPDPTTGTGLRINFRSKDMDEVDVRLPFOAEMLTQNIIVLFFCVGMAGVFPPELVAVGP	102
Db	960	KFPDPTTGTGLRINFRSKDMDEVDVRLPFOAEMLTQNIIVLFFCVGMAGVFPPELVAVGP	101
QY	1021	LVILFSLVIAHYSRVILRELKRLNDITQSPFLSHITSSIOGLATIHAYNKOEFLLHRYOEL	108
Db	1020	LVILFSLVIAHYSRVILRELKRLNDITQSPFLSHITSSIOGLATIHAYNKOEFLLHRYOEL	107
QY	1081	LDDNQAPPELFTCCAMRLAVRLDLISALITTTGLMIVLHNGQIPRAYAGLAISYAVOLT	114
Db	1080	LDDNQAPPELFTCCAMRLAVRLDLISALITTTGLMIVLHNGQIPRAYAGLAISYAVOLT	113
QY	1141	GLPQFVYRLASEMPARTSEVERINHHYIKTSLSEAPARIKKAPSPDPOEGEYTFENAE	120
Db	1140	GLPQFVYRLASEMPARTSEVERINHHYIKTSLSEAPARIKKAPSPDPOEGEYTFENAE	119
QY	1201	RYRRENPLVJLAKRVFTTKPKREKIGIVGRTSGKSSLGMALEFRLVELSGGCIKIDVARI	126
Db	1200	RYRRENPLVJLAKRVFTTKPKREKIGIVGRTSGKSSLGMALEFRLVELSGGCIKIDVARI	125
QY	1261	IGLADLSKSLITISIOEPQEVRLFSGVATSNLADPPNOYTEPOIMDALEPFRMKRCINOLPKLE	132
Db	1260	IGLADLSKSLITISIOEPQEVRLFSGVATSNLADPPNOYTEPOIMDALEPFRMKRCINOLPKLE	131
QY	1321	SEVANGDNFESVGEROLLICJIALRLRHCKILLDEEATPAADTETDLLIOETIREAFADCT	138
Db	1320	SEVANGDNFESVGEROLLICJIALRLRHCKILLDEEATPAADTETDLLIOETIREAFADCT	137
QY	1381	MLTIAHRLHYVLGSDRIMVLAAGGVVEFPDPSVLLSNDSSRFYAMFAAENKYAVRG	1437
Db	1380	MLTIAHRLHYVLGSDRIMVLAAGGVVEFPDPSVLLSNDSSRFYAMFAAENKYAVRG	1436

Query Match	Best Local Similarity	93.8%	Score 6855.5	DB 1	Length 1436
Matches 1354	Conservative 31	Mismatches 51	Indels 1	Gaps 1	
QY 1 MKDIDIGKEYIIPSPGYSVRERTSGTHNRREDKFRRTPLBCODALETARAEGIS 60	1 MKDIDIGKEYIIPSPGYSVRERTSGTHNRREDKFRRTPLBCODALETARAEGIS 60				
DB 1 MKDIDIGKEYIIPSPGYSVRERTSGTHNRREDKFRRTPLBCODALETARAEGIS 60	1 MKDIDIGKEYIIPSPGYSVRERTSGTHNRREDKFRRTPLBCODALETARAEGIS 60				
QY 61 LDASHMSQLRIIDDEHPKGYKHNGISALKPIPTCKHQPVDNAGLFCSMTRSWISLAR 120	61 LDASHMSQLRIIDDEHPKGYKHNGISALKPIPTCKHQPVDNAGLFCSMTRSWISLAR 120				
DB 61 LDASHMSQLRIIDDEHPKGYKHNGISALKPIPTCKHQPVDNAGLFCSMTRSWISLAR 120	61 LDASHMSQLRIIDDEHPKGYKHNGISALKPIPTCKHQPVDNAGLFCSMTRSWISLAR 120				
QY 121 VAKKGGELSMEDWYSLSKHSSDVMCRRLERIMOEELNMGVDASLRVYVIFCRTL 180	121 VAKKGGELSMEDWYSLSKHSSDVMCRRLERIMOEELNMGVDASLRVYVIFCRTL 180				
DB 121 VAKKGGELSMEDWYSLSKHSSDVMCRRLERIMOEELNMGVDASLRVYVIFCRTL 180	121 VAKKGGELSMEDWYSLSKHSSDVMCRRLERIMOEELNMGVDASLRVYVIFCRTL 180				
QY 181 ISTVCLMITQIAGFSGPAPMVVKKHLELYTQATESNLQYSLVLVGLLTETIVRSWISLATW 240	181 ISTVCLMITQIAGFSGPAPMVVKKHLELYTQATESNLQYSLVLVGLLTETIVRSWISLATW 240				
DB 181 ISTVCLMITQIAGFSGPAPMVVKKHLELYTQATESNLQYSLVLVGLLTETIVRSWISLATW 240	181 ISTVCLMITQIAGFSGPAPMVVKKHLELYTQATESNLQYSLVLVGLLTETIVRSWISLATW 240				
QY 241 AINVTGYRLRGAILITMAFKITLKLNKIKESLGEIINCSNDGQRMFPAAVGSIILAGG 300	241 AINVTGYRLRGAILITMAFKITLKLNKIKESLGEIINCSNDGQRMFPAAVGSIILAGG 300				
DB 241 AINVTGYRLRGAILITMAFKITLKLNKIKESLGEIINCSNDGQRMFPAAVGSIILAGG 300	241 AINVTGYRLRGAILITMAFKITLKLNKIKESLGEIINCSNDGQRMFPAAVGSIILAGG 300				
QY 301 PVVALLGMVYVYIIIGPGEGLGSAVFILFPYPMVAFASRLTAFFRRKCVATDERPOKME 360	301 PVVALLGMVYVYIIIGPGEGLGSAVFILFPYPMVAFASRLTAFFRRKCVATDERPOKME 360				
DB 301 PVVALLGMVYVYIIIGPGEGLGSAVFILFPYPMVAFASRLTAFFRRKCVATDERPOKME 360	301 PVVALLGMVYVYIIIGPGEGLGSAVFILFPYPMVAFASRLTAFFRRKCVATDERPOKME 360				

```

OY 361 VLTYYIKFMVWAKAFSSQVOKIREERRILEKAGYFOSITVGVAVIYVIVASVTSV 420
DB 361 VLTYYIKFMVWAKAFSSQVOKIREERRILEKAGYFOSITVGVAVIYVIVASVTSV 420
OY 421 HMTLGFDTLAAQAFVTVVFNSTFALKVTPSVKSLSEASVAVDRFKSLFLMEEVHMK 480
DB 421 HMTLGFDTLAAQAFVTVVFNSTFALKVTPSVKSLSEASVAVDRFKSLFLMEEVHMK 480
OY 481 NKPAHPHIEKKNATLWMDSSHSSIONSPKLTPPKMKDKRASRGKKEKVRLOLTHHOA 540
DB 481 NKPAHPHIEKKNATLWMDSSHSSIONSPKLTPPKMKDKRASRGKKEKVRLOLTHHOA 540
OY 541 VLAEBKGHLLDSDRSPSEEGKHGHRLORTLHSDLEIOEGLVIGCSVSGS 600
DB 541 VLAEBKGHLLDSDRSPSEEGKHGHRLORTLHSDLEIOEGLVIGCSVSGS 600
OY 601 KTSLSAIIIGMTLEGSIAISGTFAYVAQAAMILNATLNDLILFGEKDEERYNSVLS 660
DB 601 KTSLSAIIIGMTLEGSIAISGTFAYVAQAAMILNATLNDLILFGEKDEERYNSVLS 660
OY 661 CCLRPDLALPSSDLTEIGERGANSOGORISLARALYSRPSYIILDDPSALDAHVG 720
DB 661 CCLRPDLALPSSDLTEIGERGANSOGORISLARALYSRPSYIILDDPSALDAHVG 720
OY 721 NHIFSAIRKHLKSTVLFVTHOLOYLVDCEVIFPKKECITERGHEELMINDYATI 780
DB 721 NHIFSAIRKHLKSTVLFVTHOLOYLVDCEVIFPKKECITERGHEELMINDYATI 780
OY 781 FNNLLGEPPEVINSKSKTSQKODKPGRTSIRKKEKAVPREGOLVLEEKQGS 840
DB 781 FNNLLGEPPEVINSKSKTSQKODKPGRTSIRKKEKAVPREGOLVLEEKQGS 840
OY 841 VWSYVYVYIOAGGPAFLVLMALFNLVNGSTAWMLSWIKOGSNTYTKRNEFS 900
DB 841 VWSYVYVYIOAGGPAFLVLMALFNLVNGSTAWMLSWIKOGSNTYTKRNEFS 900
OY 901 VDSKMDNPNHMOYASIVLMAVMILKAIRGVYVFKGTLSASSRLHDELPRILRSM 960
DB 901 VDSKMDNPNHMOYASIVLMAVMILKAIRGVYVFKGTLSASSRLHDELPRILRSM 960
OY 961 KFEEDTPTGRILNRSKMDDEVNLRFOAEMFQNVILVFCVGMIAGVFWFLVAVP 1020
DB 961 KFEEDTPTGRILNRSKMDDEVNLRFOAEMFQNVILVFCVGMIAGVFWFLVAVP 1020
OY 1021 LVTLPSVLHIVSRVLRLEKRLDNTQSPFLSHITSSIOGLATIAHYNKGOEFLHVOEL 1080
DB 1021 LVTLPSVLHIVSRVLRLEKRLDNTQSPFLSHITSSIOGLATIAHYNKGOEFLHVOEL 1080
OY 1081 LDDNAPPEFLFCAMRWMLAVRLDISALITTTGMLIVLMHGOIPRAYGLAISVAVOLT 1140
DB 1081 LDDNAPPEFLFCAMRWMLAVRLDISALITTTGMLIVLMHGOIPRAYGLAISVAVOLT 1140
OY 1141 GLEFOTVRLASEATEAFSTVERINHYIKTSLSEAPARINKKASPDMOEGEVTFENAM 1200
DB 1141 GLEFOTVRLASEATEAFSTVERINHYIKTSLSEAPARINKKASPDMOEGEVTFENAM 1200
OY 1201 RRRENPLVLAKVSPFKIRKEKIGYGRGSGKSSIGMALFRLVELSGGCIKIDGRISD 1260
DB 1201 RRRENPLVLAKVSPFKIRKEKIGYGRGSGKSSIGMALFRLVELSGGCIKIDGRISD 1260
OY 1261 IGLADRSKLSIIPOEVLFGSTVRSNLPNOYTEDQIMDALERTHMEKCIQOLPLKLE 1320
DB 1261 IGLADRSKLSIIPOEVLFGSTVRSNLPNOYTEDQIMDALERTHMEKCIQOLPLKLE 1320
OY 1321 SEVMENGDNFVGERQLLICIARALLHCKILILDEATAAMDTEDDLIOETIREAPADCT 1380
DB 1321 SEVMENGDNFVGERQLLICIARALLHCKILILDEATAAMDTEDDLIOETIREAPADCT 1380
OY 1380 MULTIAHRLHTVIGSDRIWVLAOGVVEFDTPSVLLSNDSSRFYAMPAAMENKVAVVG 1436
DB 1380 MULTIAHRLHTVIGSDRIWVLAOGVVEFDTPSVLLSNDSSRFYAMPAAMENKVAVVG 1436

```

```

RESULT 4
MRP2_RAT STANDARD: PRT: 1541 AA.
AC MRP2_RAT 063145;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE CANALICULAR MULTISPECIFIC ORGANIC ANION TRANSPORTER 1 (MULTIDRUG
DE RESISTANCE-ASSOCIATED PROTEIN 2) (CANALICULAR MULTIDRUG RESISTANCE
DE PROTEIN).
GN ABCC2 OR CMOAT OR MRP2 OR CMP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=MISTAR; TISSUE=Liver;
RC MEDLINE=96180672; PubMed=859091;
RA Paulusma C.C., Bosma P.J., Zaman G.J.R., Bakker C.T.M., Otter M.,
RA Schaffer G.L., Scheper R.J., Borst P., Oude Elferink R.P.J.;
RT "Congenital jaundice in rats with a mutation in a multidrug
RT resistance-associated protein gene.";
RL Science 271:1126-1128(1996).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=MISTAR; TISSUE=Liver;
RC MEDLINE=96279006; PubMed=8662992;
RA Buechler M., Koenig J., Brom M., Kartenbeck J., Spring H., Horie T.,
RA Keppler D.;
RT "cDNA cloning of the hepatocyte canalicular isoform of the multidrug
RT resistance protein, CMP, reveals a novel conjugate export pump
RT deficient in hyperbilirubinemic mutant rats.";
RL J. Biol. Chem. 271:15091-15098(1996).
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RA Ito K., Suzuki H., Hirohashi T., Kume K., Shimizu T., Sugiyama Y.;
RT "Expression of the putative ATP-binding cassette region, homologous to
RT that in multidrug resistance associated protein (MRP), is hereditarily
RT defective in Bial hyperbilirubinemic rats (EHRB).";
RL Int. Hepatol. Commun. 292:292-299(1996).
CC -1- FUNCTION: MEDIATES HEPATOBLILIARY EXCRETION OF NUMEROUS ORGANIC
CC ANIONS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: MAINLY EXPRESSED IN THE LIVER.
CC -1- DISEASE: DEFECTS IN MRP2 ARE A CAUSE OF HEREDITARY CONJUGATED
CC HYPERBILIRUBINEMIA (EHRB).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MRP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL: I49379; AAC42087.1; -
DB EMBL: X96393; CA65257.1; -
DB EMBL: D86086; BAA13016.1; -
DR HSSP: P13569; INED.
DR InterPro: IPR001140; -
DR InterPro: IPR001617; -
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProSITE: PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat.
KW DOMAIN 1 26
KW DOMAIN 27 47
KW DOMAIN 48 67
KW DOMAIN 68 88
KW TRANSMEM 68 88
KW 2 (BY SIMILARITY).

```


FT DOMAIN 89 92 EXTRACELLULAR (BY SIMILARITY).
FT TRANSSEM 93 113 3 (BY SIMILARITY).
FT DOMAIN 114 125 CYTOPLASMIC (BY SIMILARITY).
FT TRANSSEM 126 146 4 (BY SIMILARITY).
FT DOMAIN 147 164 EXTRACELLULAR (BY SIMILARITY).
FT TRANSSEM 165 185 5 (BY SIMILARITY).
FT DOMAIN 186 309 CYTOPLASMIC (BY SIMILARITY).
FT TRANSSEM 310 330 6 (BY SIMILARITY).
FT DOMAIN 331 356 EXTRACELLULAR (BY SIMILARITY).
FT TRANSSEM 357 377 7 (BY SIMILARITY).
FT DOMAIN 378 433 CYTOPLASMIC (BY SIMILARITY).
FT TRANSSEM 434 454 8 (BY SIMILARITY).
FT DOMAIN 455 457 EXTRACELLULAR (BY SIMILARITY).
FT TRANSSEM 478 478 9 (BY SIMILARITY).
FT DOMAIN 479 540 CYTOPLASMIC (BY SIMILARITY).
FT TRANSSEM 541 561 10 (BY SIMILARITY).
FT DOMAIN 562 583 EXTRACELLULAR (BY SIMILARITY).
FT TRANSSEM 584 604 11 (BY SIMILARITY).
FT DOMAIN 605 967 CYTOPLASMIC (BY SIMILARITY).
FT TRANSSEM 968 988 12 (BY SIMILARITY).
FT DOMAIN 989 1029 EXTRACELLULAR (BY SIMILARITY).
FT TRANSSEM 1030 1050 13 (BY SIMILARITY).
FT DOMAIN 1051 1093 CYTOPLASMIC (BY SIMILARITY).
FT TRANSSEM 1094 1114 14 (BY SIMILARITY).
FT DOMAIN 1115 1115 EXTRACELLULAR (BY SIMILARITY).
FT TRANSSEM 1116 1136 15 (BY SIMILARITY).
FT DOMAIN 1137 1207 CYTOPLASMIC (BY SIMILARITY).
FT TRANSSEM 1208 1228 16 (BY SIMILARITY).
FT DOMAIN 1229 1230 EXTRACELLULAR (BY SIMILARITY).
FT TRANSSEM 1231 1251 17 (BY SIMILARITY).
FT DOMAIN 1252 1541 CYTOPLASMIC (BY SIMILARITY).
FT NP_BIND 667 674 ATP (POTENTIAL).
FT NP_BIND 1330 1337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1007 1007 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1010 1010 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 420 420 M -> V (IN REF. 3).
SQ SEQUENCE 1541 AA; 173583 MW; D5FB5571BFDB39 CRC64;

Query Match 28.7% Score 2100.5; DB 1; Length 1541;
Best Local Similarity 34.4%; Pred. No. 8.6e-129;
Matches 497; Conservative 262; Mismatches 476; Indels 209; Gaps 26;

QY 98 QHPYDNAGLFSCMTFSWLSLARYAHKKELSMEDVSL-----SKREES----- 142
DB 191 QTPSVTASFLSSITFSWDRIVLKGKHK-PLTLEVDYMDIDSGFKRSTYSKFEAMTDL 249
QY 143 -----DVCORLER-----LMQELNEVGPDASLRVVF-- 174
DB 250 QKARQAFQRRLQKSORKEPRTLHGINKKOSODVLVEEAKKSEKTKDYPKSWMLIKS 309
QY 175 -----CRRLILSIYCMITOLAGFSGPAFMKHLLEYQAESNIOYLLVLGLLLEI 230
DB 310 LKFFHHVILKSFLLKHLHLLVLELNPQ-LKLKLGKVSNSYWEFYCAIIMFAVTL 368
QY 231 VRSSSLATFMALNRTGVRGLRGAILTMAFKKILKLNK--KEKSLGELINICSDGGRMF 288
DB 369 IQSCDIOYFQHCYVGLGCVFTYMSSTYKAKLLLSNARQYITIGETVNLMSVDSQKLM 428
QY 289 EAAVAGSLLAGPVVALIGMTYVNIIGPTGFLGSAVFLEYPAMMEASRLTAYFRKCV 348
DB 429 DATYMWOLWMSYQIITLISFLFELRELGPSTILAGVAVLLIPVGVATLAKIRNIQVNM 488
QY 349 AADPERYOKMNEVLTLYTKITKMAVAVKAFSOSVOKIREERKILIEKAGYPOSTIVGAPI 408
DB 489 KKKKRLKIMELLISGKILKLYFAMEPESFOYOGIRKREKELKNLIRFQGLSLIFILQI 548
QY 409 VVVIASVVTESVHMTLGD--LTRAQAFVTVTVFNSTFALKVTPEFSKLSSEASVAVDR 466
DB 549 TPLIVSVVTFESVYVLYVSANVLANEKAFTSTITLNLIRFLPSMLPMVTSTILOASVSVD 608

QY 467 EKSLDMEVHMKKPPASPHK-IEKNRATLAWDSHSSIONSKLIPKMKKKRASRG 525
DB 609 IERYLGGDLDTSAIRVSNFDKAVKFSSEASPTWD----- 643
QY 526 KKEKVRQLRTEHQVLAEQKHLLDSDERSPEEBEGKHILGHLRLQRTLSIDLEI 585
DB 644 -----PD-----LEATIQDNVNDI 657
QY 586 QEGKLVGICGSVSGKTSLSAIILOMTLLEGSIAISGFVAVAOAMILNATLNDNLF 645
DB 658 KPGQVAVAVGVGSSKSLVSLAMLGEMENVHGHITIGSTAVVPQSMONGTIDNILE 717
QY 646 GKEVEDEERNVLANCCRLPDALIPSSDLTEIGRGANLSCGQQRISLARALYSDSI 705
DB 718 GSEYNEKKYQVLKACALLPDLEILPGGDMAIEGKEKINLSGCKORVSLARAAYQADI 777
QY 706 YLLDDPLSLADHVNHNHFNENSAIRKH--LKSUTVAFVHQLOLYVDCEVIFMKGCITE 763
DB 778 YLLDDPLSLADHVNHNHFNENSAIRKH--LKSUTVAFVHQLOLYVDCEVIFMKGCITE 837
QY 764 RGTHEIMNLNGDYATIFNNLLGETPVEI----- 794
DB 838 KGSYRDLDDKKGVAFARNNKTFPKHSGPGEAEVYVNDSEAEDDDLITMERIPEDASL 897
QY 795 -----NSKRET-SGSQKKSQDKP-----KTGSIKKRAVKEGQVQLEKKGGS 840
DB 898 AMRRENSLRRTLSRSSRSSRSGKSLKNSLKIKNVNLKEKEKEVEGGKLIKKEVEYERK 957
QY 841 VPWSYGYVYIQAGPLAFVYLMALFNLVNGSTASTWMLSVYWGSGNTVTYGNETS 900
DB 958 VKFSYLYLYLQAVG-WSLFLLIFLYGLNNAFVIGSNLMSMTSD--SDNT-----NQTN 1010
QY 901 VSDSKDNPDMQYASIVASIALMAVLLIKALRGVYFVKGTL-----RASRLDELEFR 953
DB 1011 NSSSHD-----MRGVGALGAGIOLITLSTLMSIYACRNASALHGOULT 1058
QY 954 RILRSPMKFEDTPTGRLILNRESKMDDEVYVLPQAEMLQNVILVEFCVG---MAG 1009
DB 1059 NLRAPMFEFTPTPGRIYVNRSGDISYVDLLP-----QTLRSMWMCFFIAGTLVIMCM 1114
QY 1010 VEPWFLVAVGPVILFSLVHLVSRVILREIKRDNITQSPFLSHITSSIOGLATHAANK 1069
DB 1115 ATPVAILITIPLSLITISVQVYVATSRQLRLDSVTSPISHSSEVTGIPITIRAEH 1174
QY 1070 GQEFRLRYQELDDNQAFEFLETCAMRLAVRLDISIALITTTGLMTVLMHGOIPRAYA 1129
DB 1175 QQRFLANNEKQDINDQKCVFMSITNSRMLAIRLEVLGNLVVFCSSALLVIYIKTITGIV 1234
QY 1130 GLAISYAVOLTGLFOFVRLASSETAEFTSYERINHYIKTISLEAPARIKKKAPSPMPQ 1189
DB 1235 GFVLSNALNITQTLNWLVRMSEAEITNIVAEIRISEYI--VNEENAP--WYTDKRPADMPR 1292
QY 1190 EGEVTFEAMERRENDLVLYKKVSFTIKPKKIGIVERTSSGKSLCMALFRLVELSGG 1249
DB 1293 HGEIOFNHYQVTRPDLVLKNGITCNKSGEKGVAVRTAGKSLINCLFRILLESAGG 1352
QY 1250 CIKIDVRIISIDGLADRSKSIIPQEPVLFSGVRSMLDFPNQTEQIDALERTMKN 1309
DB 1353 QIIDGIDIVASIGLHDLERLITIPQDPILESGLSRMLDPPNKSDSEVWRALELAHR 1412
QY 1310 ECIQOLPLKLESEVWENDNFSVGERQLLCLARALLRCKILLIDEATAAMDTEITDILQ 1369
DB 1413 SFVSGLOGLLSEVTEGSDNLSIGORQLLCLGRAVLRKSKILLIDEATAAVDLEDSILO 1472
QY 1370 ETIRAFADCTMTLTAHLHLVLSGDRIMVLAQGVAFEDIPSVLLSNDSSRFAMF--A 1427
DB 1473 TTIRKFSQCTVITTAHLRHTIMDSKIMVLDNKKIYEGSPBEELSNRGS--FYIMAKEA 1531
QY 1428 AAKN 1431
DB 1532 GLEN 1535

RESULT 5
MRP1_HUMAN STANDARD; PRT: 1531 AA.
ID MRP1_HUMAN P33527; O14819; F78419;
AC 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MULTIDRUG RESISTANCE-ASSOCIATED PROTEIN 1.
GN ABCB1 OR MRP OR MRP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9308080; PubMed=1360704;
RA Cole S.P.C., Bhargava G., Gerlach J.H., Mackie J.E., Grant C.E.,
Almquist K.C., Stewart A.J., Kurz E.U., Duncan A.M.V., Deeley R.G.;
RT "Overexpression of a transporter gene in a multidrug-resistant human
lung cancer cell line." Science 258:1650-1654 (1992).
RL Science 258:1650-1654 (1992).
RN [2]
RP REVISIONS.
RX MEDLINE=93262415; PubMed=8098549;
RA Cole S.P.C., Deeley R.G.;
RT "Multidrug resistance-associated protein: sequence correction." Science 260:879-879 (1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96251691; PubMed=8649356;
RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M.,
Cole S.P.C., Deeley R.G.;
RT "Structure and expression of the messenger RNA encoding the murine
multidrug resistance protein, an ATP-binding cassette transporter." Mol. Pharmacol. 49:962-971 (1996).
RN [4]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=98008927; PubMed=9344662;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
resistance protein gene (MRP) and alternative splicing of its mRNA." Genomics 45:368-378 (1997).
RN [5]
RP SEQUENCE OF 1131-1531 FROM N.A.
RX MEDLINE=99425270; PubMed=10493829;
RA Loftus B.J., Kim U.-J., Sneeddon V.P., Kalush F., Brandon R.,
Fuhmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from
human chromosome 16p and 10q." Genomics 60:295-308 (1999).
RN [6]
RP TOPOLOGY, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=97442425; PubMed=9295302;
RA Hifner D.R., Almquist K.C., Leslie E.M., Gerlach J.H., Grant C.E.,
Deeley R.G., Cole S.P.C.;
RT "Membrane topology of the multidrug resistance protein (MRP). A study
of glycosylation-site mutants reveals an extracytosolic NH2
terminus." J. Biol. Chem. 272:23623-23630 (1997).
RN [7]
RP TOPOLOGY.
RX MEDLINE=97476249; PubMed=9334225;
RA Kast C., Gros P.;
RT "Topology mapping of the amino-terminal half of multidrug resistance-
associated protein by epitope insertion and immunofluorescence." J. Biol. Chem. 272:26479-26487 (1997).
RN [8]
RP TOPOLOGY.
RX MEDLINE=98153110; PubMed=9485377;
RA Kast C., Gros P.;
RT "Epitope insertion favors a six transmembrane domain model for the

RT carboxy-terminal portion of the multidrug resistance-associated
RT protein." Biochemistry 37:2305-2313 (1998).
RL BIOFUNCTION: MAY PARTICIPATE DIRECTLY IN THE ACTIVE TRANSPORT OF
CC DRUGS INTO SUBCELLULAR ORGANELLES OR INFLUENCE DRUG DISTRIBUTION
CC INDIRECTLY.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: A NUMBER OF ISOFORMS ARE PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: LUNG, TESTIS, AND PERIPHERAL BLOOD
CC MONONUCLEAR CELLS.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MRP SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L05628; AAB46616.1; -;
CC EMBL; AF022853; AAB83983.1; -;
CC EMBL; AF022824; AAB83983.1; JOINED.
CC EMBL; AF022825; AAB83983.1; JOINED.
CC EMBL; AF022826; AAB83983.1; JOINED.
CC EMBL; AF022827; AAB83983.1; JOINED.
CC EMBL; AF022828; AAB83983.1; JOINED.
CC EMBL; AF022829; AAB83983.1; JOINED.
CC EMBL; AF022830; AAB83983.1; JOINED.
CC EMBL; AF022831; AAB83983.1; JOINED.
CC EMBL; AF022832; AAB83983.1; JOINED.
CC EMBL; AF022833; AAB83983.1; JOINED.
CC EMBL; AF022834; AAB83983.1; JOINED.
CC EMBL; AF022835; AAB83983.1; JOINED.
CC EMBL; AF022836; AAB83983.1; JOINED.
CC EMBL; AF022837; AAB83983.1; JOINED.
CC EMBL; AF022838; AAB83983.1; JOINED.
CC EMBL; AF022839; AAB83983.1; JOINED.
CC EMBL; AF022840; AAB83983.1; JOINED.
CC EMBL; AF022841; AAB83983.1; JOINED.
CC EMBL; AF022842; AAB83983.1; JOINED.
CC EMBL; AF022843; AAB83983.1; JOINED.
CC EMBL; AF022844; AAB83983.1; JOINED.
CC EMBL; AF022845; AAB83983.1; JOINED.
CC EMBL; AF022846; AAB83983.1; JOINED.
CC EMBL; AF022847; AAB83983.1; JOINED.
CC EMBL; AF022848; AAB83983.1; JOINED.
CC EMBL; AF022849; AAB83983.1; JOINED.
CC EMBL; AF022850; AAB83983.1; JOINED.
CC EMBL; AF022851; AAB83983.1; JOINED.
CC EMBL; U91318; AAC15784.1; -;
CC PIR; A44231; DVHVAR.
CC HSSP; P13569; INBD.
CC MIM; 156343; -;
CC InterPro: IPR001140; -;
CC InterPro: IPR001617; -;
CC Pfam: PF00664; ABC membrane; 2.
CC Pfam: PF00005; ABC tran; 2.
CC PROSITE: PS00211; ABC TRANSPORTER; 2.
CC ATP-binding, glycoprotein; Transmembrane; Transport; Repeat;
CC Alternative splicing.
CC DOMAIN 1 33
CC TRANSMEM 34 54
CC DOMAIN 55 74
CC TRANSMEM 75 95
CC DOMAIN 96 100
CC TRANSMEM 101 121
CC DOMAIN 122 133
CC TRANSMEM 134 154
CC DOMAIN 155 172
CC TRANSMEM 173 193

```

FT DOMAIN 194 316 CYTOPLASMIC.
FT TRANSMEM 317 337 6.
FT DOMAIN 338 363 EXTRACELLULAR.
FT TRANSMEM 364 384 7.
FT DOMAIN 385 440 CYTOPLASMIC.
FT TRANSMEM 441 461 8.
FT DOMAIN 462 464 EXTRACELLULAR.
FT TRANSMEM 465 485 9.
FT DOMAIN 486 547 CYTOPLASMIC.
FT TRANSMEM 548 568 10.
FT DOMAIN 569 590 EXTRACELLULAR.
FT TRANSMEM 591 611 11.
FT DOMAIN 612 967 CYTOPLASMIC.
FT TRANSMEM 968 988 12.
FT DOMAIN 989 1025 EXTRACELLULAR.
FT TRANSMEM 1026 1046 13.
FT DOMAIN 1047 1089 CYTOPLASMIC.
FT TRANSMEM 1090 1110 14.
FT DOMAIN 1111 1113 EXTRACELLULAR.
FT TRANSMEM 1112 1132 15.
FT DOMAIN 1133 1203 CYTOPLASMIC.
FT TRANSMEM 1204 1224 16.
FT DOMAIN 1225 1226 EXTRACELLULAR.
FT TRANSMEM 1227 1247 17.
FT DOMAIN 1248 1531 CYTOPLASMIC.
FT NP_BIND 678 685 ATP (POTENTIAL).
FT NP_BIND 1327 1334 ATP (POTENTIAL).
FT CARBOHYD 19 19 N-LINKED (GLCNAC. . .).
FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .).
FT CARBOHYD 1006 1006 N-LINKED (GLCNAC. . .).
FT VARSPPLIC 706 764 MISSING (IN ISOFORM WITHOUT EXON 17).
FT VARSPPLIC 765 820 MISSING (IN ISOFORM WITHOUT EXON 18).
FT VARSPPLIC 1431 1495 MISSING (IN ISOFORM WITHOUT EXON 30).
SQ SEQUENCE 1531 AA; 171560 MW; 5CCB92FEB7B096B52 CRC64;

```

Query Match 28.5%; Score 2082; DB 1; Length 1531;

Best Local Similarity 34.5%; Pred. No. 1.4e-127; Matches 488; Conservative 262; Mismatches 479; Indels 184; Gaps 21;

```

100 PVDNAGLESCMTFSLSSLAIVAHKKGELSMEDVWSLSKSHSDVNCBRLRLMOEELNE 159
Db 209 PESSASPLSRITFTFWITGLI-VRGYROPLEBSDLINKEDTSEGVVYLYKMKKECAK 267
Qy 160 V-----GPDA-----SLRRVWIFCFRLIL 181
Db 268 TRKQPVKVVYSSKDPAQKESKVDANEVEALIVKSPQKEMNPSLKFVLYKTEGPFILM 327
Qy 182 STVCLMITOLAGFSGPAMVWHLEIYQATSNLQYSLLVGLLTITIVSMGLATWA 241
Db 328 SEFFKAIHDMFMFSGPO-ILKLILKEVNDTRKAPMVGFFYVLLFVTAQLQTLVLYHOYFH 386
Qy 242 LNVTRGVLRGAIITLMAKKTILKLNKERS--LGEIINISNDGORMFEAAVAGSLIAG 299
Db 387 ICFPSGRKIKTAIVGAVYRAKRLVITNSARKSSTYGEIYNLMSVDAORPMDLATYINMTWS 446
Qy 300 GPVVAALGMIYVNIILGTFGLSGLAVFLFYPAAMFASRLTAFFRKCAVATDERVQKMS 359
Db 447 APLOVLIALLYLLMNLGPSVLAVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 506
Qy 360 EVLYYIFIKKAYAVKAFASQVOKIREERKILKAGFQSIYGVAVIYVAVIYVAVIYVAV 419
Db 507 EILNGIKVLYLAMELFEKDVLAIRQELVLYKSAVLSAVGFTTWCTPELVALCTFA 566
Qy 420 VHMTLGPD--LTAQAFVTVVNSMTFALKVTPPSVKSLSAEASVAVDRFSLMEEVH 477
Db 567 VYVIDERNIILDAQAFVSLFLNLRPLNLRPLNLRPLNLRPLNLRPLNLRPLNLRPLNLRPL 626
Qy 478 --MIKNRPASP--HIKTEMNATLAMDSSHSISONSPLKPKMKKDRASRGKKEKVRQ 532
Db 627 PDSIERRRVVKGGGNSITVNAFTW----- 653
Qy 533 LQRTBHOAVLAEQGHLLDSDERPSPPEEKGHIHLGLRLQRTLHSLIDLEIQRGLVIG 592

```

```

Db 654 -----ARSDP-----TINCITSIPEGALVA 675
Qy 593 ICGVSGKTSLLSAILGQMTLLEGSLAISGTFAYVAQOAILNATLNDILFKEXEDY 652
Db 676 VVGVOVGCKSSLLSALLAEMDKVEGHVAKISGVAVVPOAWIOWDSIRENITLFCQLEEP 735
Qy 653 RYNSVNSCCRLPDLATIPSSDLTEIGRGANLSGGORQRISLARALYSDSIYIILDDPL 712
Db 736 YRYSYIQAALLPDLLEILPSGDRTEIGRGVNLGGQKORVSLARAYSNADYILFDDPL 795
Qy 713 SALDAHVNHIENSAT--RKHLKSTVLFVTHOLOYVDCDEVIFMEKGCITERTGHEEL 770
Db 796 SAVDAHVCKHIFENVIGRKGLKTKRILTYHMSYLPQVNVIIYMSGKISEMGQVEL 855
Qy 771 MNLNGDYATIF-----NNLL-----GETPPEIN 795
Db 856 LARDAFAEFLRTYASTRQODAEENGVTGVSGPKKAKQKMGNCMLVDSAGKOLQROL 915
Qy 796 SKKETSGQKKSQDKGPKTSGIKKEKAVKPEEGOLYOLEKGGQSVPSYGVYIQAAG 855
Db 916 SSSYSYSGDISRHHN--STAELQKAERKEETWKLMEADKQGTQVKLSYWDYWKALGL 972
Qy 856 PLAFVLIALFNLVNGSPAFSTWMLSYIKQSGNTVYTRGNETSVSDMKDNPHMOYVA 915
Db 973 FISTFSLIF-LFMCHNVSLAANYMLSLMTDDPIVNGT--QEHKKVRLSYVGAIGISGI 1028
Qy 916 STYALSMVAMLLKAIKGVFVKGTLRASSRLHDELFRILLRSPKKEFTPTPTGRIILNF 975
Db 1029 AVFGYMAVSI-----GGIILASRCHLDLISILRSPSPFEPFSGNLVRF 1076
Qy 976 SKMDDEVYRLPQAEKMTQVAVLYFPCVGMAGVFPFELVAVGPIVLYFSLVHIVSVL 1035
Db 1077 SKELDTVDSMIPEYIKMKWGLFVNIAGACIVILATPAAIILPPLGLTY--FFVQREY 1133
Qy 1036 I--RELKRLNITQSPRLSHITSIOGLATTHAVNKGCEFLHRYOELDDNDQAPFLFT 1092
Db 1134 VASSQQLKRLSESVSNPSYSHFNELTGLSVYIRAFEBERITHQSDLKVDENQAYYISI 1193
Qy 1093 CAMRLAVRLDLISALITTTGLMIVLMHGOIPRAYAGLAISYAVOLTGLOFTVRLASE 1152
Db 1194 VANRLAVRLBOVCNCIYLAFLFAVIRSHISLAGVLGLSVSYHQVTTYVNLVLMVRMSE 1253
Qy 1153 TEARTSVIRINHYIKTILSLAPARIKKAAPSPPMPQGEYTFEAKERYENRPLVLYK 1212
Db 1254 METNIVAEERLKEYSET-EKAPWQIOETAPSSWPQVGRVFRNYCYLRVREDLDFVLRH 1312
Qy 1213 VSEFIKPKKXIGYVGRSGSSGLMALFRLVELSGGCIKIDGVIRSDIGLADLRSLKSI 1272
Db 1313 INVITNGEKVGIYGRTAGSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTI 1372
Qy 1273 IPOEVLFGSTVRSNLDPFNQYTEDQIDALERTHMKECIAQLPLKLESEVMENGDFSV 1332
Db 1373 IPQDVLFGSGLRMNLDPFSQYSDDEVMTSLELAHLKDFVSALPDKLHDECAEGENISV 1432
Qy 1333 GEROLCTIARALLRCKILLIDETAAAMDTEYDILLQITIEARFADCTMLTARHLTVL 1392
Db 1433 GQROLYVCLARALLRCKTKLLIVDEAFAAVDLEFDDLLQSTIRTFQEDCTVLTARLNTIM 1492
Qy 1393 GSDRIMVLAOGVVEFDPVSVLSDSSRYAM 1425
Db 1493 DTYRIVLDKGEIETYGAPSDLL-QQRLGFTSM 1524

```

RESULT 6

MRP2_HUMAN STANDARD: PRT: 1545 AA.
 ID MRP2_HUMAN 092887; 092963; 092987; 014022; 092500;
 AC 092887; 092963; 092987; 014022; 092500;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE CANALICULAR MULTISPECIFIC ORGANIC ANION TRANSPORTER 1 (MULTIDRUG
 RESISTANCE-ASSOCIATED PROTEIN 2) (CANALICULAR MULTIDRUG RESISTANCE

Query Match	Best Local Similarity	Score	DB 1:	Length	1325;
Match 489: Conservative	255;	Mismatches 465;	Indels 171;	Gaps 33	
FT	TRANSMEM	351	371	POTENTIAL.	
FT	TRANSMEM	440	460	POTENTIAL.	
FT	TRANSMEM	710	730	POTENTIAL.	
FT	TRANSMEM	771	791	POTENTIAL.	
FT	TRANSMEM	836	856	POTENTIAL.	
FT	TRANSMEM	858	878	POTENTIAL.	
FT	TRANSMEM	954	974	POTENTIAL.	
FT	TRANSMEM	977	997	POTENTIAL.	
FT	TRANSMEM	1038	1058	POTENTIAL.	
FT	NP_BIND	445	452	ATP (POTENTIAL).	
FT	NP_BIND	1075	1082	ATP (POTENTIAL).	
FT	CARBOHYD	651	651	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	690	690	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	746	746	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	754	754	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	792	792	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	1176	1176	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	1309	1309	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CONFLICT	1302	1302	H -> D (IN REF. 2).	
SEQ	SEQUENCE	1325	AA: 149539	MM: 9C5750A748BB96CE	CRC64:

Qy	99	HPVONAGFCSOMTFSWMLSLARVNHKGGELSMEDWMSLSKHESSDVNCRRLERJMOEJLN	158
Db	11	NPLODANICSVFVFWMLNPLFKIGHKR-RLBEDMYSVLEPDSOHJGELOGFWDKYVL	69
Qy	159	EVGDA--ASLRVWVIFCTRLLISVCLMITOLAGSGSPAFMVNKHLEYTO-----	209
Db	70	RAENDADKPSLTRAIIKICYMWSYLVLEIFTL-IEESAKVIOPIFJLG-IINFEENDPMD	127
Qy	210	--ATESNOYSLLAVIGLLILEIYRMSLSLTMALNVR-----TGVRLGAILMAPFKIL	263
Db	128	SVALNTAAYATVITFECTLI-----LAILHLYTHYQACAGMRJLRAMCHMIRKAL	179
Qy	264	KLKNIK--EKSLGELINICSDGOMEAAVGSLLAGPRVVALGMIVYVILGPTFL	321
Db	180	RLSNMAGKTTGGIVNLSNDVKRFQVYVFLHFLWAGPLQALAVATALLMEIGISCLA	239
Qy	322	GSAPFLLFYPAMASULTAFVRKCYAADDEROKNNEVLYTKFKTMAMWKAPSOY	361
Db	240	GMAVAILLPLQCSFGKLFSLRSKRTATFTDARKITNEVITGIRITIKAMESESNTLI	299
Qy	382	OKIREE-----RILEKAGYFOSITGVAPIVYVAVSVTFESVHTLGFDDLTAQ	432
Db	300	TNLKKEIKSLIRSSCJLRGNMLASFSA-----SKIIV-----PYTFYTYLVISVITASR	350
Qy	433	AFYVYVFNSTPLKAYT--PESVKSLSSEASVANDRFKSLFIMEEYHMITKN--PASPH	487
Db	351	VFVAVTTLGAV--SLYTWLFFPSPALIEVSEALVISIRIQFLLDLDEISO--RNRQLSDDK	407
Qy	488	IKIKENKATTLAMDSSHSIONSPLPLFRMKKKDKKAKSGKKENKQOLRTEHQAVALDOKG	547
Db	408	KMAYVODTAWDK-----	421
Qy	548	HLLLDDEPSPSEEBEKKHHLGLRLQRTLSIDLEIOGKVLGVSOGTSLSIA	607
Db	422	-----ASBTP-----TLQGISFVVRGELLAAVVGPAAGKSSLSLA	457
Qy	608	ILGOMTLLEGSIALSGTFAYVAAQAMLTALRNILFGKYEDEERKNASVUNSCCLRPDL	667
Db	458	VLGELAPSHGIVASHGRVAAVSODPWFSGTLRNLILFGKRYEKERKAYIKACALIKDL	517
Qy	668	AALPSSDLTEIGERGANLISGOGORISIALRYSIDRSIYLLDEPLSLADHVNHNIPNSA	727
Db	518	QMLEBGDLTVIGDRGTLTSSGQAKRVMLNARVYDADYIYLLDPLSLAVDAEVSHNLELC	577
Qy	728	IRKHLKSTVLEVTHOLOYLVDCEVIFMKEGCTITERGTHEELMNLNGVATYATFNNLL-	766
Db	578	ICOLIHKEITVLVHOLQVLAASOILIKKGGKVAOKGVTAEFLKSGID-----FGSLTK	633

Oy	787	--GEPFPPV----	EINSKKETIGSOKKQDKPKTGSIKKKAVKPREGGDVL-----E	834
Db	634	DNESEDPVPVGPPTLNNRFSSESSVWSOOSSRPSL---	KDGALSDOTENVPVTLSEE	689
Oy	835	EKGGSVPMSVYGVYIOAAGCPPLAFIVIALMFLANGST---	AFTSWMLSMVI-KOCSGN	890
Db	690	NRSCKVKGFQATKNYPAGA--HAIYVFILLINTAAOVAVI	LVDWMLSTRANKOSMLN	746
Oy	891	TTVV-RGENETSVSDSMKMDPNMOYAYISALSMAVALILKAIRGVV	FVKGTLRASKRRLHD	949
Db	747	VTVNGGDNVTEKLD-----LMWYIGIYSGLTVAITVLEGIARSLIV	PYLVYNSSQTLLHN	799
Oy	950	ELFRILRASPKREDDTPPGRIILNRSCXMDENVDRLPPOAQAMFI	ONIIYLFPCWGIMAG	1009
Db	800	KMFPSILKAPVLEPFDNRPIGRILNRRSKDIGHLDDLPLETIDF	IQTLLQYGVAVSAVA	859
Oy	1010	VFPFPIVAVGPLVLTLSFSLIHSRVILIRELKRLDNTIOSPELSHTSS	IOGLATIHAVNK	1069
Db	860	VIPMIALPVPPLGIITFLFLARVFLERSRDVKRLRESTRTSPVS	SHLSSSLQGIMTRAY-K	918
Oy	1070	GGEFLHRYOEILLDNO---APPFLFTCAMRLAVRLDLISTALITTT	GLMTVLMGOIP	1125
Db	919	AEE---RCOEELFDAHQDLSEAMFLELTTSRMFAVBLADICAMFYI	IVAFGSLIAKTLD	975
Oy	1126	PAYAGLIASVAYOLTGIFOSTVLAESTARTSVSRINHYYIKTSL	EPARIKKKASP	1185
Db	976	AGGYGLASTLVTLMGMFOVCQKHOSEVEEMMISERVLET-TDLEKAP	MPEYO-KRBP	1033
Oy	1186	DMPOEGEVTEFNAMERREMPLVLTKKVSFTIKPKKIGIVSGTSGKS	SJGMLFLIVE	1245
Db	1034	AMPHEGVILIDNVNFMKSPGGPPLVTLKLHALIKSQKGVIVERTG	AGASSLISALFRE	1093
Oy	1246	LSGCCIKIDGVRISDGLADLSKLSIIPOEPLVLSGIVRSULDPNP	OTYTEDQIDALER	1305
Db	1094	PEGK-IWIDKITLTEIGLHDBRKMSIIPQEPVLTGTMRKMLDPKE	RTEBELMALQOE	1152
Oy	1306	THMKECIAOLPLKLESEVMVNGDFNSVGEROLLCJARILRCKLI	LIDEATAAMDTETD	1365
Db	1153	VQLMETIEDLPGRKMDTLASSGNSFVGOROLVCLARALIRKNOILI	IDETANVDPRTD	1212
Oy	1366	LLIOTIEREAFADCTMLTIARHLRTVLSGDRIWYAQQGVFEEDP	SPVLLSNSDSRFYAM	1425
Db	1213	ELIORKIREKPAHCTVLTIAHRILTIIIDSPKIWLDSGRKFKYDE	PVYLONKESELFYKM	1272
<hr/>				
RESULT	8			
MRP2_RABIT	ID	MRP2_RABIT	STANDARD;	PRT; 1564 AA.
AC	028689;			
DT	30-MAY-2000 (Rel. 39,	Last sequence update)		
DT	30-MAY-2000 (Rel. 39,	Last annotation update)		
DE	CANALICULAR MULTISPECIFIC ORGANIC ANION TRANSPORTER 1 (MOLTIDRUG	RESISTANCE-ASSOCIATED PROTEIN 2) (CANALICULAR MULTIDRUG RESISTANCE		
DE	PROTEIN) (EPITHELIAL BASOLATERAL CHLORIDE CONDUCTANCE REGULATOR).			
GN	ABC2 OR MRP2 OR BECR			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_Taxid=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-Ileum.			
RX	MEDLINE=96224297; PubMed=8643587;			
RA	van Klijck M.A., van Aubel R.A.M.H., Busch A.E., Lang F.,			
RA	Russel F.G.M., Bindels R.J.M., van Os C.H., Deen P.M.T.;			
RT	"Molecular cloning and expression of a cyclic AMP-activated chloride			
RT	conductance regulator: a novel ATP-binding cassette transporter.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 93:5401-5406(1996).			
RN	[2]			
RP	FUNCTION.			
RX	MEDLINE=98279125; PubMed=9614209;			


```

DB 1112 IMICATPFPAVILIPILAIIVYAVGVATVATSNQRLRLDSVFRSPYISHTETEVSLPVI 1171
QY 1065 HAYNKGQELHRYOELLDDNOAFELFTCAMRLAVRLDLISALTITGLMIVLMHGOI 1124
DB 1172 RAPEHQORFLKONEIGIDYNOKCVSSMITSNRLAPRLVGLNMFVSSALMVIYRDLT 1231
QY 1125 PRAYGLAISYANOLGLQFYVRLASFEARTSVERTINHYIKTISLEPAIKKAPB 1184
DB 1232 SGDVGVFVLSNANITQTNLMVLRMTSETEINIVAEVTEYTK-VENEXP-WVTXKRP 1289
QY 1185 PMPQEGEYFENAEKRYENRPLVLYKRVSTFTPEKKEIGIVRGSGKSSLMGLFRLY 1244
DB 1290 AGMPHNGEIGFNSYQVRYRPELDVLKGINCDIKSMKIGVORTAGKSSLLNCFRL 1349
QY 1245 ELISGCIKIDGVRIISDGLADLRKSLITIPQEVLFSGVYRSLMDPPNOYTEDQIDALE 1304
DB 1350 EAAGCHITIDGIDIASIGLHDLRGKLTITIPDVFSGSLRMLDPPNNYSDEEIRALE 1409
QY 1305 RTHMKECIAQLPLKLESEVMENGDNFSGVEROLLCTARALLRCKLIIIDEATAADTET 1364
DB 1410 LAMKSFVAGLHGLSREVSSEADNLISGROLCLGRALLRCKLIVLDEATAADLET 1469
QY 1365 DLIIDETIEAFADCTMLTIAHRLATVLSGSDRIMVLAQGVPEPTPSVLSN 1417
DB 1470 DHLIQTINNEFSHCVTITIAHRLHTIMDSKIMVLDNGNIVYSGPELLES 1522

```

```

RESULT 9
YCFI_YEAST STANDARD: PRT: 1515 AA.
AC P39109: 003905;
DT 01-FEB-1995 (Rel. 31, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE METAL RESISTANCE PROTEIN YCFI (YEAST CADMIUM FACTOR 1).
GN YCFI OR YDRI35C OR YD9302.11C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H9;
RA OLIVER K., HARRIS D., BARRELL B.G., RAJANDREAM M.A.;
RT Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RL J. Biol. Chem. 269:22853-22857(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA OLIVER K., HARRIS D., BARRELL B.G., RAJANDREAM M.A.;
RT Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RL [3]
RP FUNCTION
RA MEDLINE=20253522; PubMed=10790694;
RA Petrovic S., Pascolo L., Gallo R., Cupelli F., Ostrow J.D.,
RA Goffeau A., Tiribelli C., Bruschi C.V.;
RA "The products of YCFI and YDL015W (BPT1) cooperate for the ATP-
RT dependent vacuolar transport of unconjugated bilirubin in
RT saccharomyces cerevisiae.";
RL Yeast 16:561-571(2000).
CC -1- FUNCTION: COOPERATES FOR THE ATP-DEPENDENT VACUOLAR TRANSPORT OF
CC BILIRUBIN.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MRP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC EMBL: L35237; AAA50353.1; -
CC EMBL: 248179; CAA88217.1; -
DR HSSP: P13569; INRD.
DR SGD: S0002542; YCFI.
DR InterPro: IPR001140; -
DR InterPro: IPR001617; -
DR Pfam: PF00664; ABC_membrane; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transmembrane; Glycoprotein; Transport;
KW Cadmium resistance; Phosphorylation.
FT DOMAIN 1 32
FT TRANSMEM 33 53
FT DOMAIN 54 73
FT TRANSMEM 74 94
FT DOMAIN 95 99
FT TRANSMEM 100 120
FT DOMAIN 121 130
FT TRANSMEM 131 151
FT DOMAIN 152 169
FT TRANSMEM 170 190
FT DOMAIN 191 278
FT TRANSMEM 279 299
FT DOMAIN 300 345
FT TRANSMEM 346 366
FT DOMAIN 367 422
FT TRANSMEM 423 443
FT DOMAIN 444 466
FT TRANSMEM 467 487
FT DOMAIN 488 530
FT TRANSMEM 531 551
FT DOMAIN 552 572
FT TRANSMEM 573 593
FT DOMAIN 594 943
FT TRANSMEM 944 964
FT DOMAIN 965 1001
FT TRANSMEM 1002 1023
FT DOMAIN 1024 1066
FT TRANSMEM 1067 1087
FT DOMAIN 1088 1088
FT TRANSMEM 1089 1109
FT DOMAIN 1110 1180
FT TRANSMEM 1181 1201
FT DOMAIN 1202 1205
FT TRANSMEM 1206 1226
FT DOMAIN 1227 1515
FT NP_BIND 663 670
FT NP_BIND 663 670
FT NP_BIND 856 863
FT MOD_RES 1306 1313
FT MOD_RES 908 908
FT MUTAGEN 713 713
FT MUTAGEN 908 908
FT CONFLICT 680 680
SQ SEQUENCE 1515 AA; 171120 MW; 30F92FDBDA60431 CRC64;

```

Query Match 27.1%; Score 1983.5; DB 1; Length 1515;
 Best Local Similarity 33.1%; Pred. No. 3.6e-121;
 Matches 472; Conservative 268; Mismatches 492; Indels 195; Gaps 27;

```

QY 77 PKGKYHGLSALPIRTCKHQHPVDNAGLFCSTFWSLSLARVAHKKGELSMEQVWSL 136
DB 196 PKHHIQTILRRKP-----NPYSANIFSRITTSWMSGKLTGVEK-YLVEADYKYL 246
QY 137 SKHESSDVNCRRLERLMOEELNEVGPDASLRVYVFCPT---RLILSLVCLMTITQLAG 193
DB 247 PRNFSSEELISQKLEKRWENELKOKSPSIS-----WALCRTEFSKMLLAFFRAIDVLA 301

```

```

QY 194 FSGP---AFWKHLEXT-----GATESNLOYSLLVGLLT-----TE 229
Db 302 FTQPOLRLIKFTYDINSEKODHSSLOGFENHPKRLPYRGFLIAPAMELVGTOTS 361
QY 230 IVRSMSTALTYALNYRGVRLGALITFWAFKILKLNKE--KSIIGELINCSNDGORM 287
Db 362 VLHQYFLN-----FNGMYIKSALTALITYOKSLVLSNEAGSLSTGDIYMLMSVDYOKL 416
QY 288 FEAAVAGSLAGVVALIYVNIIVIIIPGTGSAVFILFYPAAMFASRLTAYFRKRC 347
Db 417 ODLQWMLNLSGCPOLITCLYSYLKLGNSMWGVILLVIMPLINSFLMIQOKLQKQSQ 476
QY 348 VAADERQKNNVEITYIKFTKMYAMVAFSOSQOKIR--EERRRLKAGFQOSTITVEA 406
Db 477 MKYDERTRVISELLNNKSKLYAMEKPYREKLEEVNKNELNLTGCGYMAATSFQF 536
QY 407 PIYVIVSVWFESVHM--TLGFDLTAAGAFVTVFVNSMTFALKYVPFSVKSLSASVAVD 465
Db 537 NIVPPLVSCFFFAVFVYTEDRALTTDLVFPALITLFPNLISFLMITPMVINSFTEASVIG 596
QY 466 RFSKSLFMEVYHMKRNPASPHIKEMKNATLWDSHSSIONSFKLTPKMKKKDRASRG 525
Db 597 RLFFFTNEEL-----QPDVQRLPKV--KNIGVAINIG 629
QY 526 KKEVROLOQREHOAVLAEOKGHLILSDERPSPEEKGKIHGHLRLQRTLSIDLEI 585
Db 630 DDAATFLMORKPEYKV-----ALKNNINQA 653
QY 586 OEGKLVIGCVSGSKTSLIAILGOMTLBGSIAISGTPAVYAOAMILNATLNDITLF 645
Db 654 KKGNIITCVGKSGKTALILSCMLGDLFRVKGFAIVHGSVAIVSQVPIAMGTAVENILF 713
QY 646 GKEYDEERYNSVLNSCCLRPDLAILPSSDLTEIGERANLSGGGQORISLARALYSDBSI 705
Db 714 GHRDAEYKTIKACALITDILAIMDQDKTLVEKGISLISGQCARISLARAVYARADT 773
QY 706 YILDDPLSALAHGNNHFNFSAIRKH--LKSQTVLFVTHOQLYVDCBEVFMKGCCTE 763
Db 774 YLLDDPLAAVEHVARHLEHVLGPNGLHTKTVLATNKSALSIADSIALLDGEITG 833
QY 764 RGTHEEL-MNNGDVFATFNNLLIGETPPVEINSKETSOGSKSODKPGPKGISKKEKA 822
Db 834 QGTDEITKDDSPMLKLLNNY-----GKKNNG--KSNERGDSSESSVRESS 878
QY 823 VKPEGQULVOL-----EKGQGSVPMS 844
Db 879 I-PVGELEQLOKUNDLDFGNSDAISLRASDATLGSIDGDDENIARKEHQKVKWN 937
QY 845 VYGVYIOAAGPIAFVLVMALEFNLVNSTAFSTWLSYWIQSGSNTVTYKGNETSVD 904
Db 938 IYLEVAKAC-NKRSYCVFILFIVISMFLSYGVGNVWLKHM-----SEVNSR 981
QY 905 MKDNPMMOYVASYI---ALMAVYMLILKAIGVYVFKGTLRASRLHDELFRRIIRSPK 961
Db 982 YGSNNNARYALAIYALDGLSALTLIOTI--VLWVECTIASHKSLHMLNMTSVYARAPT 1039
QY 962 FPDTPTRIGRIINRESKMDVEDVRLPFOAEMFIQNVILVFCVGMIAGVFPWFELVAGPL 1021
Db 1040 FFEETTPIRIIRFENDIYKADALLGRFSQOFVNAVAVTFTTIVICATWQFIILIP 1099
QY 1022 VILFVLAHIVSRVLIREKRLDNTQSPFLSHITSIOGLATIHAYNKGQELHNYOELL 1081
Db 1100 SVEFYVYQOYLRFSRERLRDSTIRSPYSHFOETLGLAVRAGYSOQKRSHNOCRT 1159
QY 1082 DDNOAPFLFTCAMRLAVRLDLISALITTTGLMIV--LMHGQPPPYAGIAISYAVOL 1139
Db 1160 DNNSAFYPSINANRWLALYRLIELISIIILGAALISVRLKGTITLAAVAGLSLSYALQI 1219
QY 1140 TGLFOFVRLASETEARTSVYRINHYTKTISLEAPARIKKNKASPDMPQSEVTFENAE 1199
Db 1220 TOTLNMIVRMTEVETNISVERIKY--ADLKSEAPLIVEGRHPKPEKPSQDIFKNYNS 1278
QY 1200 MRYRENPLVLVKKVSFTIKPKKIGIVRTSGSKSLMALFLRLVLESGGCKIKIDGVRIS 1259

```

```

Db 1279 TRYREEDLVKHNINHKPKENKGIYVGTGAGKSSLLALFPMIEASEGNIVIDNIAIN 1338
QY 1260 DIGLADSKSLIIPQEVLFSGVYRSMDDPNQYTEDQINDALERTHMKACIAQLPLK- 1318
Db 1339 EIGLYDLRHKLSIIPQDSQVEGVRENIIDPINOYTDIAIRALELSHKHVSMSMDG 1398
QY 1319 LSEVPMENDPVSVERDLCLARALLHCKIILITDEATAAMDQETDILLOETIREFAD 1378
Db 1399 LDAOITTEGGMLSVGQROLCLLARMLVPSKILVLEDAVTAADVETDKVVOETIRTAFKD 1458
QY 1379 CMTLTAIRLHTVLGSDRIWYLAOGOVEFDPYSVLISNDSSRYAM 1425
Db 1459 RTIILTAIRLHTIMDSRITIVLDNKAFAFDSPOGLLSDNKSIFYSL 1505

RESULT 10
MRP3_RAT STANDARD: PRT: 1522 AA.
ID MRP3_RAT 088563;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CANALICULAR MULTISPECIFIC ORGANIC ANION TRANSPORTER 2 (MULTIDRUG
DE RESISTANCE-ASSOCIATED PROTEIN 3) (MRP-LIKE PROTEIN-2) (MLP-2).
GN ABCC3 OR CMOAT2 OR MRP3 OR MLP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=99292429; PubMed=10362653;
RA Ortiz D.F., Li S., Iyer R., Zhang X., Novikoff P., Arias I.M.;
RT "MRP3, a new ATP-binding cassette protein localized to the canalicular
RT domain of the hepatocyte."
RL Am. J. Physiol. 276:G1493-G1500(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Colon;
RX MEDLINE=98279126; PubMed=9614210;
RA Hitchashi T., Suzuki H., Ito K., Ogawa K., Kume K., Shimizu T.,
RA Sugiyama Y.;
RT "Hepatic expression of multidrug resistance-associated protein-1 like
RT proteins maintained in eisel hyperbilirubinemic rats."
RL Mol. Pharmacol. 53:1068-1075(1998).
CC -1- FUNCTION: MAY ACT AS AN INDUCIBLE TRANSPORTER IN THE BILIARY AND
CC INTERSTITIAL EXCRETION OF ORGANIC ANIONS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: LUNG AND INTESTINE, LOW IN LIVER. HIGHER IN
CC LIVER OF EISAI HYPERBILIRUBINEMIC RATS AND TR(-) MUTANT RATS.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MRP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL: AF072816; AAC25416.1; -
CC EMBL: AB010467; BAA28955.1; -
CC HSSP: P13569; INBD.
CC InterPro: IPR001140; -
CC InterPro: IPR001617; -
CC Pfam: PF00664; ABC_Membrane; 2.
CC Pfam: PF00005; ABC_Tran; 2.
CC PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat.
FT DOMAIN 1 32 EXTRACELLULAR (BY SIMILARITY).

```

FT	TRANSMEM	33	53	1 (BY SIMILARITY)
FT	DOMAIN	54	73	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM	74	94	2 (BY SIMILARITY).
FT	DOMAIN	95	99	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM	100	120	3 (BY SIMILARITY).
FT	DOMAIN	121	132	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM	133	153	4 (BY SIMILARITY).
FT	DOMAIN	154	171	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM	172	192	5 (BY SIMILARITY).
FT	DOMAIN	193	301	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM	302	322	6 (BY SIMILARITY).
FT	DOMAIN	323	347	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM	348	368	7 (BY SIMILARITY).
FT	DOMAIN	369	424	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM	425	445	8 (BY SIMILARITY).
FT	DOMAIN	446	448	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM	449	469	9 (BY SIMILARITY).
FT	DOMAIN	470	531	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM	532	552	10 (BY SIMILARITY).
FT	DOMAIN	553	574	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM	575	595	11 (BY SIMILARITY).
FT	DOMAIN	596	958	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM	959	979	12 (BY SIMILARITY).
FT	DOMAIN	980	1016	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM	1017	1037	13 (BY SIMILARITY).
FT	DOMAIN	1038	1080	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM	1081	1101	14 (BY SIMILARITY).
FT	DOMAIN	1102	1123	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM	1103	1123	15 (BY SIMILARITY).
FT	DOMAIN	1124	1194	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM	1195	1215	16 (BY SIMILARITY).
FT	DOMAIN	1216	1217	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM	1218	1238	17 (BY SIMILARITY).
FT	DOMAIN	1239	1522	CYTOPLASMIC (BY SIMILARITY).
FT	NP_BIND	659	666	ATP (POTENTIAL).
FT	NP_BIND	1318	1325	ATP (POTENTIAL).
FT	CARBOHYD	18	18	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1001	1001	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1002	1002	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	323	344	SSSSSHSCSSSSGLFRPHGY -> LSFINDPQLSLILRFR
FT	CONFLICT			ISDPNAPT (IN REF. 2).
FT	CONFLICT	645	645	I -> L (IN REF. 2).
FT	CONFLICT	1075	1075	H -> D (IN REF. 2).
Q0	SEQUENCE	1522 AA:	169977 MM:	440E31E0C4664297 CRC64;

[illegible]

QY	410	VVLSVSVTPESHMHLGCD--LTLAQAFTVYVWPNXMTFLKTPSPVXSLSBASVANDR	467
Db	541	PPWVTLITLGYVCVDKNKNVLADEKAFVSLFENLKIPLNLIPOLISGMOTYSILKRI	600
QY	468	KSLFLMEEV--HMKNKPASPPIKIKEMAKNATLAMPSSHSIONSPLKPKKKDKORASRG	525
Db	601	QDPLNODLDPQCYERKTIISPGRALITIHNGTFSW-----SKD-----	637
QY	526	KKEKVRQJORTHOAVLAEQGHLLDSDERSPEEHEKHIHLGLRLORTLSIDLEI	585
Db	638	-----LPLTHSINIQI	649
QY	586	QEGKLVGICGSVSGKTSLSIALIGOMTLLEBSIAISGTFPAVQAQAMNLATLNDLNF	645
Db	650	PKGALIVAVGVCGGKSLVSALLGEMEKLEGASVGSVAVPOQAMIONCTLOENLFL	709
QY	646	GKEVDERKYNVLSNCCLRDLALIPSSDLTEIGEAGALSGGQORISIALRALSDSI	705
Db	710	GQPMNPKRYQOAEFTCLALDLDPVLPBGQOTETEGEKIILSGQORQSLAAVSDNI	769
QY	706	YTLDDPSALDAHAGNIFINSAIRKH--LKSQVLEFVTHQLOYLVDCEVIFMKEGCITE	763
Db	770	FLDDPLSADSHAKIIPPOVIGPEGVLAGKTRVLYTHGISFLQOTDPIVLADGQITE	829
QY	764	RGTHEELMNLNGDYATIFNN-----LLIGET-----PP	791
Db	830	MGHVSELLQHDGSPANFLRYNVPADQENAEVGLQANEBVLLLEDTLSTHDTLDTPEA	889
QY	792	V-EIINSKETSIGSKKSDQDGR-----TGSIKKE--KAVPREGQOLULEEGQOSY	841
Db	890	IYEVKQFQFMREMSLSSEGGQNRPLKRYTSSLEKEVPATQFKETGALIKELIAETGVN	949
QY	842	PMSYGVYIOAAGGPLAFVIALMPLMNVGSTFF--SPWMLSYWIKOGSGTWTVRG--	896
Db	950	KLSIYMDYAKVG---LCTTFLTCILLYAQNAVAIGANWILSANT-----NDVEHQO	1000
QY	897	NETSVDSMKDNPMOYTYASIALMAVMILKAIKRGVVFVKGTLRASRLHDELFRIL	956
Db	1001	NNTSV-----RLGYAYTL-GILGGLIYMLSAFTMV---GAJOARLHTALLHNOI	1048
QY	957	RSPKPFEDTPTTGTILNRSEKDEVDYRLPFOAEFIONVILV-----FCYGMAG	1009
Db	1049	RAPOSFEDPTPSGRILRFSKDIYIH-----EVLAPFTILMFNSEFTYSISTIVYA	1101
QY	1010	VFPMPVAVGVLVLFVSUHLHVSRLIRELKRDLTOSPLSHITSSIOGATIHAYNK	1069
Db	1102	STPLECVVLPFAEYFQVPRFYVATSRQIKRLESYRSRPIFSHSEYVGTGIVIRAGR	1161
QY	1070	GQEFLLHRYOELLDNQAPFLFTCAMKMLAVRLDLISALLITTTGIMLVIMHGOIPAYA	1129
Db	1162	VQDFPVLSDAVDSONQTTYPIATASNMGVHNHEFYGNCVYLFASLPAVIGRNSINPGLV	1221
QY	1130	GLAISYAVQLTGLFQFTVRLASETEARFTSVSEINHYITKLSLEAPARIK--NKAPSPPM	1188
Db	1222	GLISYSYALQVTLSTNMWIRLPLSDESNITAVEVKEYSKT--ETEPAPWLESNRAPE--GMP	1279
QY	1189	QEGEVTPEFNAMARREMLPYLTKKVSPTIKPKKEIGIVORTSGGSSGMALFRLVYELSG	1248
Db	1280	RSGVEEFENNSVRRPGELVEVNLKNTLHVQGGKGVAVGRTGAGSSMKLFRILEAAE	1339
QY	1249	GCIKIDVTRISDIDIALDRKSLSIIPQEPVLFEGYRSNIDPFNOYTEDQIDALERTHM	1308
Db	1340	GEIFLDGLNVAHIGLHDLRSQTLTIPODPIFLFSGTILRMULDPEGRXSDPDIWRTLESHL	1399
QY	1309	KECTIAQPLKLESVEWNGNFSYGEKOLICARALLRCKIILIDDEATAMDTDLII	1368
Db	1400	SAFVSQPTGLDFQCSGCGNLSYGOLOICLARALKRSRYLVDEATAADLETDDLI	1459
QY	1369	QETIRREAPDCTMTLTIAHRLHTYVGSRLIMVLAQOVVEFDPSPVSLSDSSRFAYM	1425
Db	1460	QGTIRTOGEDCTVTLTAHRLNTIMDKIRVYLVADKGVAAEDSVNLTA--AGSIGTGM	1515

Qy	1010	VEFVFIIVAGGVILFESVLIHIVSRVILIREKRLDINTOSPFISHTISSQIGLATIHAANK	1069
Db	1102	STPLECVVILFELAFYEFVQRFYVATVSROKRLIESRSRPISHSFSEVITGVSVIARA	1161
Qy	1070	GOEFELHRYOELLDDNQAPFELFTCAMRWLAVRDLISALITTTGMLIVIMHQIIPAYA	1129
Db	1162	VQDFRVLSDAKVDSNQNTTPTPIASNNWIGVHFEFGNCVVFELFALFVIGNSLNPLV	1221
Qy	1130	GLAISAVOLTGLEFQFTVRLASSETEARFVSVERINHYIKTLSLEAPARIK-NKAPSPMP	1188
Db	1222	GLSVSYALQVTLSTNMMLIRLSDLESNTIAVEKKEYSKT-ETEPAPWLESNRAPE-GMP	1279
Qy	1189	QGEVTEFPMNARVRREMLPIVLIKVSPTIKKREKIGVORTSGGSSGMALFRLVEISG	1248
Db	1280	RSGVVEFNVSVRPQGLEIVLKNLTLHVQGVKGVIIVERTAGVSSMILCFRLLEAAE	1339
Qy	1249	GCIRKIDVRSIDIGIADLRKSLSIPOEPVLFSGVRSNMDPFNOYTEDIWDALERTHM	1308
Db	1340	GEIFLDGLNVANHIGLHDLRSQTLTIIPQDPILEFSGTIRMLMDPFGCRSDPIDWRLTLESHL	1399
Qy	1309	KECIAQDPLKLESSEVMENGSFVSERGEOLICARALLRCKLILIDETAAQDTDLII	1368
Db	1400	SAFVSQPTGLDFQCSGEGCNLSVGGOICLARALLRSRYLVDEATAADILETDDLI	1459
Qy	1369	QETIRREAPDCTMTLTIAHRLHTVLSGRIVLVAGVGGVFEFDPSPVSLTSDSSRFAPAM	1425
Db	1460	QGTIRTOGEDCTVTLTAHRLNTIMDKVIRVLVDKGVAAEDSVNLTA-AGSIFEGM	1515

RESULT 11
MRP3_HUMAN STANDARD; PRT; 1527 AA.
ID MRP3_HUMAN 060265; 060922; 075621; 095078; 095289; 095290;
AC 015438; 060265; 060922; 075621; 095078; 095289; 095290;
DT 15-JUN-1998 (Rel. 36, Created)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE CANALICULAR MULTISPECIFIC ORGANIC ANION TRANSPORTER 2 (MULTIDRUG
DE RESISTANCE-ASSOCIATED PROTEIN 3).
OS ABC3 OR cMOAT2 OR MRP3 OR MRP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC NCBI_TaxID=9606;
RN MRP3_HUMAN
RP SEQUENCE FROM N.A.
RX MEDLINE=98409330; PubMed=9738950;
RA Kluchl Y., Suzuki H., Hirohashi T., Tyson C.A., Sugiyama Y.;
RT "cDNA cloning and inducible expression of human multidrug resistance
RT associated protein 3 (MRP3).";
RL FEBS Lett. 433:149-152(1998).
RN MRP3_HUMAN
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Keppeler D., Koenig J.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN MRP3_HUMAN
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=99032812; PubMed=9813153;
RA Uchilund T., Hinooshita E., Haga S., Nakamura T., Tanaka T., Toh S.,
RA Furukawa M., Kawabe T., Wada M., Kagotani K., Okumura K., Kohno K.,
RA Akiyama S., Kuwano M.;
RT "Isolation of a novel human canalicular multispecific organic anion
RT transporter, cMOAT2/MRP3, and its expression in cisplatin-resistant
RT cancer cells with decreased ATP-dependent drug transport.";
RL Blocham. Biophys. Res. Commun. 252:103-110(1998).
RN MRP3_HUMAN
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Kool M., de Haas M., Ponne N.J., Baas F., Borst P.;
RT "Complete coding sequence of human MRP3, a homolog of the human
RT multidrug resistance-associated protein MRP1.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN MRP3_HUMAN
RP SEQUENCE FROM N.A. (ISOFORMS 3A AND 3B).
RC TISSUE=Liver;
RX MEDLINE=99107643; PubMed=9889399;
RA Fromm M.F., Leake B., Roden D.M., Wilkinson G.R., Kim R.B.;
RT "Human MRP3 transporter: identification of the 5'-flanking region,
RT genomic organization and alternative splice variants.";
RL Blocham. Biophys. Acta 1415:369-374(1999).
RN MRP3_HUMAN
RP SEQUENCE OF 1043-1527 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97413640; PubMed=9270026;
RA Kool M., de Haas M., Scheffer G.L., Scheper R.J., van Eljk M.J.,
RA Juljn J.A., Baas F., Borst P.;
RT "Analysis of expression of cMOAT (MRP2), MRP3, MRP4, and MRP5,
RT homologues of the multidrug resistance-associated protein gene
RT (MRP1), in human cancer cell lines.";
RL Cancer Res. 57:3537-3547(1997).
RN MRP3_HUMAN
RP FUNCTION: MAY ACT AS AN INDUCIBLE TRANSPORTER IN THE BILIARY AND
CC INTESTINAL EXCRETION OF ORGANIC ANIONS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 3 (SHOWN HERE), 3A AND
CC 3B; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: MAINLY EXPRESSED IN THE LIVER. ALSO EXPRESSED
CC IN SMALL INTESTINE, COLON, PROSTATE, TESTIS AND BRAIN.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MRP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AB010867; BAA28146.1; -
DR EMBL; Y17151; CAAT6658.1; -
DR EMBL; AF083552; AAC34668.1; -
DR EMBL; AF009670; AAD01430.1; -
DR EMBL; AF085690; AAD02845.1; -
DR EMBL; AF085691; AAD02846.1; -
DR EMBL; AF083692; AAD02847.1; -
DR EMBL; U83659; AAB71756.1; -
DR HSSP; P13569; INBD.
DR MIM; 604323; -
DR InterPro; IPR001140; -
DR InterPro; IPR001617; -
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Alternative splicing.
FT DOMAIN 1 32
FT TRANSMEM 33 53
FT DOMAIN 54 73
FT TRANSMEM 74 94
FT DOMAIN 95 99
FT TRANSMEM 100 120
FT DOMAIN 121 132
FT TRANSMEM 133 153
FT DOMAIN 154 171
FT TRANSMEM 172 192
FT DOMAIN 193 302
FT TRANSMEM 303 323
FT DOMAIN 324 349
FT TRANSMEM 350 370
FT DOMAIN 371 426
FT TRANSMEM 427 447
FT DOMAIN 448 450
FT TRANSMEM 451 471
FT DOMAIN 472 533
FT TRANSMEM 534 554
FT DOMAIN 555 576
FT TRANSMEM 577 597
FT DOMAIN 598 963
FT TRANSMEM 964 984
FT DOMAIN 985 1021
FT TRANSMEM 1022 1042
FT DOMAIN 1043 1085
FT TRANSMEM 1086 1106
FT DOMAIN 1107 1107
FT TRANSMEM 1108 1128
FT DOMAIN 1129 1199
FT TRANSMEM 1200 1220
FT DOMAIN 1221 1222
FT TRANSMEM 1223 1243
FT DOMAIN 1244 1527
FT NP_BIND 661 668
FT NP_BIND 1323 1330
FT CARBOHYD 18 1006
FT CARBOHYD 1007 1007
FT VASPLIC 226 510

CC EXTRACELLULAR (BY SIMILARITY).
CC 1 (BY SIMILARITY).
CC CYTOPLASMIC (BY SIMILARITY).
CC 2 (BY SIMILARITY).
CC EXTRACELLULAR (BY SIMILARITY).
CC 3 (BY SIMILARITY).
CC CYTOPLASMIC (BY SIMILARITY).
CC 4 (BY SIMILARITY).
CC EXTRACELLULAR (BY SIMILARITY).
CC 5 (BY SIMILARITY).
CC CYTOPLASMIC (BY SIMILARITY).
CC 6 (BY SIMILARITY).
CC EXTRACELLULAR (BY SIMILARITY).
CC 7 (BY SIMILARITY).
CC CYTOPLASMIC (BY SIMILARITY).
CC 8 (BY SIMILARITY).
CC EXTRACELLULAR (BY SIMILARITY).
CC 9 (BY SIMILARITY).
CC CYTOPLASMIC (BY SIMILARITY).
CC 10 (BY SIMILARITY).
CC EXTRACELLULAR (BY SIMILARITY).
CC 11 (BY SIMILARITY).
CC CYTOPLASMIC (BY SIMILARITY).
CC 12 (BY SIMILARITY).
CC EXTRACELLULAR (BY SIMILARITY).
CC 13 (BY SIMILARITY).
CC CYTOPLASMIC (BY SIMILARITY).
CC 14 (BY SIMILARITY).
CC EXTRACELLULAR (BY SIMILARITY).
CC 15 (BY SIMILARITY).
CC CYTOPLASMIC (BY SIMILARITY).
CC 16 (BY SIMILARITY).
CC EXTRACELLULAR (BY SIMILARITY).
CC 17 (BY SIMILARITY).
CC CYTOPLASMIC (BY SIMILARITY).
CC ATP (POTENTIAL).
CC ATP (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC MAITYGRHPLPEKRLMSKEDRSOMVVOOLEAMROEKO
CC TARRKASABEKNSGDEVILGARRPRRSFKALIAAT
CC GSSFLISACEFLIDLDLFINPQLSLIRFISNPAKSMW
CC GFLVAGLMFLCSMQSLILQHYHYIVTGKFTGIMGYI
CC YRKALVITNSYKRASTGVEIVNLMSVDAORMDLAPLNL
CC WSAPIQITAIYFLMONIGPVLAVGAVMLIPLNGAVAY
CC KMRAPVOKMKIKDNRISIKMSEITINGIKVILKYIWMDSF
CC -> LINDPLRGCLPGFTSPDGHMLPASGGSGPLVPKG

FT RGQIPDGGAAAAAGMEAGKADGTTQGFSTWCKLRBGRG
FT AAGCPAPAEALPPEGPAGHLRLDLPOLHCLQAYGPPALLH
FT OSTAQHDPDYOYLPHGHLGLPBGMDNVLPHAVADLT
FT TLPLHCLDWEVSTWHDGCHLQBSGSHQUSOTVHCNGN
FT CQPHVSGCPALHPCPLQASAVSTPADHGGDLPLAPRS
FT LCPGMSRPHGLADSTORSGREDARLPKANEIEGLAHQAD
FT E (IN ISOFORM 3B).
FT MISSING (IN ISOFORM 3B).
FT VARSPLIC 511 1527
FT VARSPLIC 1194 1527
FT OVTFLANMIRMSDLESINIAVERKETSLETAPWVE
FT GSRPEGPMPKGEVEFRNYSVRPGLDVLNIDSLHVG
FT EKVGIKRTGAGSSMTCLFRLLEAAGEIIDELNAD
FT GLHDLRSQITIPDPIFSGTLRMLNLPFGSSEEDIMWA
FT LELSHLTFVSSOPAGIDPFCGEGENISVGOROIYCLARA
FT LLKRSITIVDEATATIDLETNLTQITRQFDCTYVIT
FT AHRNTIMDYTRVLDGVAAKDFNSPAILNARIFIGMA
FT RDAGLA -> SEASALPCCSSRNQOALWCSSISLSPKO
FT KTGAPLPLPHFLI (IN ISOFORM 3A).
FT C -> R (IN REF. 2).
FT FHGLVHGR -> SMASMA (IN REF. 2).
FT A -> T (IN REF. 5).
FT A -> G (IN REF. 3).
FT MELCSM -> DVPVAPW (IN REF. 4).
FT F -> Y (IN REF. 1).
FT F -> C (IN REF. 6).
FT L -> I (IN REF. 6).
FT D -> E (IN REF. 6).
FT CLF -> ACS (IN REF. 6).
FT L -> F (IN REF. 6).
FT LRSOL -> VRCOM (IN REF. 6).
FT L -> V (IN REF. 3).
FT L -> R (IN REF. 6).
FT Q -> R (IN REF. 6).
FT SEQUENCE 1527 AA; 169342 MW; 0D1F879B6F18370C CRC64;

Query Match 26.6%; Score 1945; DB 1; Length 1527;
Best Local Similarity 32.3%; Pred. No. 1.2e-118;
Matches 464; Conservative 268; Mismatches 510; Indels 196; Gaps 22;

QY 85 LSAI-----KPIRTCK-----HOHPVDNAGLFSCMTFSWLSLARVAHKKGELSMEDV 133
DB 182 LSAIILACREKRPFPFSAKVNDPNPYPERSAGFLSRLFFWMFTKMAIYGRH-PLEEKD 240
QY 134 WLSLHSSDVNCRLEKRMQELNEVGDAS-----L 167
DB 241 WSLKEEDRSQMVVOLLLEAMROEKOTARHKASAPGNASGEDEVLLGAPRRKPSFL 300
QY 168 RRVVWIFCETRLILSLVCLMITQIAGFSGAPFMVKKHLEYQTATESNLQYSLILVGLLL 227
DB 301 KALLATF-GSSFLISACFKLIQDLDLSFINPO--LSTILIRFISNPMAPSWMGFLVAGLMFL 358
QY 228 TEIVWSMALTWALNTRTVGRVLRGAILTMAFKKILKLN-IKES-LGELINICSDCO 285
DB 359 GCMQSLILQHYHYHIFVGVKFRGIMGVIRKALVITNSVKRSTVEIYMLMSVDQ 418
QY 286 RNFEEAANGSLLAGGVVAAILCMITIVIIILPFGELGSAVFILFPAMMPASRLAYFRR 345
DB 419 RMDMLAPFNLTNLSAPLQIILAIYFLMOMIGPSVLAVFAMVLLPLNGAVVAKMRAQV 478
QY 346 KCVAAATDEVOKANEVLYIKFIKMYAWKAPSOYOKIRREERRILKAGYFQSTIVGV 405
DB 479 KQMKLKDSRIKLMSELINKIKVCLKYAMPSPFLKQVEGIRQGLRLKRYAALYHTTTTPT 538
QY 406 ---APIVVIVASVTFVSHMTLGFDTLAAQAFVTVVFNFSMTFALKVTPFSVKSLSSEASV 462
DB 539 WMCSPFLVTLITLIMYV-YVVDNNVLDAAKAFYSVSLFNLRLPLMLLQLLSNLQASV 597
QY 463 AYDRKSLFLAMEV--HMKNRPASPHIKIEKNATLANDSSHSIIONSPKLTTPKAKK 520
DB 598 SLKRIQOFLQELPOQSVYERKTIISPGVAITTHSGTFAOD----- 639
QY 521 RASRGKKEKVRLOQRTFHQAVLAQKGHLLDSDERPSPEEEKGHIHGHLLQRTLHS 580
I 1111

DB 640 -----LPTLHS 646
QY 581 IDLEIOEGKLVGICVSGSKTSLISALIGOMTLLEGSAISGTFAYVAQOAMILNATR 640
DB 647 LDIOVKGALVAVVGVGGKSSLSVALIGEMEKLEGKVMGSAVYVQOAMIOCTQD 706
QY 641 DNILFKEDEDERVNSVLSCLRPDLAIPSSDLTEIEIGERANLSCGQROISILARALY 700
DB 707 ENVLKGLANPKRYQOTLEACALLADLEMLPGDQTEIEIGKINLSGQROISILARAY 766
QY 701 SDRSIYIIDPLSALDAVGHNI FNSAIRKH--LNSKTVLFYTHOLOYLVDCEVLEKME 758
DB 767 SADIFLDDPLSADVSASHAKHIFDHVIGPEGLAGKTIVLVTHGISFLPDQDFIVLAD 826
QY 759 GCITTEGTHBELMNLNGDYA-----TIFNN 783
DB 827 GQVSEMGPPALLQKRGSPANFLCNYPAPDEQGHLEDSTALEGADKELLIEDTLSSH 886
QY 784 LLLGETPPV-----EINSKKEKTSQKSKODKPKTSGIKREKAVKPEEGOL 830
DB 887 TDLITNDPVTYYVQKQFMQSLALSDEGQGRPVYRRHLGP---SEKQVTEAKADGAL 943
QY 831 VOLLEKQGSVPMSYVYVYIOAAGPLAFLVJMALFMLNVGSTAF--STWMLSYWIKOG 887
DB 944 TQEKKAICTVELSVWDYAKAVG-----LCTTLAICLLVYGSAALIGANVWLSAW----- 995
QY 888 SCNTYTRKNETSVSDSMKDNPHMQYTAITYALSMAVMLILKAIKGVVVKGLRASSSL 947
DB 996 -----TNAMADSRONNTSLR--LGVTYALGILQGFVLWLAAMAAGIOQARVL 1044
QY 948 HDLFRILRISPMKPFDDTPGTRILNRFKSKDDEVYVRLPQOEMPTOMVILVFCVGM 1007
DB 1045 HQALHNKIKRSPQSFDTIPSGRILNCSKDIYVDEVLAIVYMLNLSFNNAISLTVYI 1104
QY 1008 AGVFPWFLVAVGPVLILFSLVHIVSRVLIRELRKLDNITQSPFLSHITSSIOGLATIHAY 1067
DB 1105 MASTPLFTVILPLAVLYTLVQRFYAATSRQLKRLSEVSRSPIYSHFSEVVGASVIRAY 1164
QY 1068 NKQGFPLHRYQELLDNQAPFLFTCAMRWLAVRLDLISALITTTGMLVLMHQIIPA 1127
DB 1165 NRSRDEIISDKVDANQSCPYIISNRMLSIGVEVNCVVLFPALPAVIGRSSLNG 1224
QY 1128 YAGLAISYAVOLTGLQFVIVRLASETEARFTSYERNHIIKTLSEAPARIKKKAPSPW 1187
DB 1225 LVGLSVSYSLQVTFALNMMIRMSDLESINI VAEKESKT-ETAPVWEGSRPEGM 1283
QY 1188 PQEGEYTFENAEKRYRENPLVLYKVSFTIKPEKIGIYGRSGKSSGLMAFLRVELS 1247
DB 1284 PRGEYEFERNYSVRYPGLDYLRDLSLVHGEKKGIVGRICAGKSSMTCLCFRLLEA 1343
QY 1248 GGCIRKIDVIRISDIGLADLRSLKSLIIPQEPVLESGTVRSNLDPEFNOYTEDQIWDALERTH 1307
DB 1344 KGEIRIDGLNADIGLHDLRSQTLTIIPDPILESGLTRMLDPFGSYSBEDIWMALELSH 1403
QY 1308 MKECIAQLPLKLESEVMEGDNFVSGBEQLLCAARLLHCKLILIDEXATAADTTDL 1367
DB 1404 LHPFVSSOPAGIDPQCSGEGENISVGOQLVCLARLLKRSRLVLDXETAADLETMDL 1463
QY 1368 IOETIEAPADCTMLTIFARLHTVLGSDRIMVLAQOQVEFDPSPVSLNDSRFYAM 1425
DB 1464 IQATITQTDCTIVLTIAIRLNTIMDYTRVLYVDKQVVAEFSPALIN-ARQIFFGM 1520
RESULT 12
YAMB_SCHPO
ID YAMB_SCHPO STANDARD; PRT: 1478 AA.
AC Q10185;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROBABLE ATP-DEPENDENT PERMEASE C3F10.11c.
GN SPAC3F10.11c.
Schizosaccharomyces pombe (Fission yeast).

DB 1102 SUGSLTIRAYDMEDFISENDIRVDNHRHWFELFYSSNRMOAIRVEAIGALVVFSSAF 1161
 OY 1117 IYL--MHQOIPRANGLAISVAVOLTGLFQFVVRCLASEFAFTVERHNHKLISLEA 1174
 DB 1162 GLSARVGNPSGLVLSLSTAVQITOSTLFVVRQSVDEINIVSERMLEYI-GLPSSA 1220
 OY 1175 PARINKKAPSPDPOEGEVTEFNAMRRYENLPVLKKKFTIKPEKIGIYRTGSGKS 1234
 DB 1221 PSTIDHRPREMPHSGAIKPFHYSVRRENPLVINDISVNIKQEKIGIYRTGAGKS 1280
 OY 1235 SIGMALFRLVELSGGCIKIDGVRSIDIGLADLRSLIIPQEPVLFGSTVRSLNLPFNOY 1294
 DB 1281 TLTFLALFRIEPTSGDIDINITSIGLHDLRSRLAITPQENQAFEGTIRRENDPNANA 1340
 OY 1295 TSDQWDLERHMECAQPLKLESEYMGNDNFVGEROLLCIARLLHCKILLID 1354
 DB 1341 TDEIWHLEASLAKQFQTLQDGGYSRTEGGANLSSQORLMCTRLTLPTRVLLID 1400
 OY 1355 EATAMDETDLIQETIREAFADCTMLTIAHRLHTVLSGDRIMVLAQGVFEPTPSVL 1414
 DB 1401 EATAVDVEITDAIVQRTIRENDRTILTIHRTVVDNSNRLVLDHGVYEDSTKTL 1460
 OY 1415 LSNDSSREFYAM 1425
 DB 1461 LENKASLFYSL 1471

RESULT 13
 ACC8_HUMAN STANDARD; PRT; 1580 AA.
 AC Q09428; 075948; Q16583;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE SULFONYLUREA RECEPTOR 1.
 GN ABC8 OR SUR1 OR SUR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gonzalez G., Aguilar-Bryan L., Bryan J.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RA Nishimura M., Miki T., Aizawa T., Seino S.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ForeSkin;
 RA Thomas P.T., Wohlik N., Huang E., Gagel R.F., Cote G.J.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1186-1580 FROM N.A.
 RC TISSUE=Pancreatic Islets;
 RA Thomas P.M., Cote G.J., Wohlik N., Haddad B., Mathew P.M., Rabi W.,
 RA Aguilar-Bryan L., Gagel R.F., Bryan J.;
 RT "Mutations in the sulfonylurea receptor gene in familial persistent
 RT hyperinsulinemic hypoglycemia of infancy."; Science 268:426-429(1995).
 RL [5]
 RP TOPOLOGY.
 RP MEDLINE=99436117; PubMed=10506167;
 RA Raab-Graham K.F., Cirillo L.J., Boettcher A.A., Radeke C.M.,
 RA Vandenberg C.A.;
 RT "Membrane topology of the amino-terminal region of the sulfonylurea
 RT receptor."; J. Biol. Chem. 274:29122-29129(1999).
 RL [6]
 RP REVIEW ON VARIANTS.

RX MEDLINE=99268411; PubMed=10338089;
 RA Meisner T., Beilbrech B., Mayatepek E.;
 RT "Congenital hyperinsulinism: molecular basis of a heterogeneous
 RT disease."; Hum. Mutat. 13:351-361(1999).
 RL [7]
 RP VARIANT PPHI VAL-715.
 RX MEDLINE=96354544; PubMed=8751851;
 RA Thomas P.M., Wohlik N., Huang E., Kuhnle U., Rabi W., Gagel R.F.,
 RA Cote G.J.;
 RT "Inactivation of the first nucleotide-binding fold of the
 RT sulfonylurea receptor, and familial persistent hyperinsulinemic
 RT hypoglycemia of infancy."; Am. J. Hum. Genet. 59:510-518(1996).
 RL [8]
 RP VARIANT ALA-1368.
 RX MEDLINE=96220088; PubMed=8635661;
 RA Inoue H., Ferrer J., Welling C.M., Elbein S.C., Hoffman M.,
 RA Mayorga R., Warren-Perry M., Zhang Y., Mills H., Turner R.,
 RA Province M., Bryan J., Permutt M.A., Aguilar-Bryan L.;
 RT "Sequence variants in the sulfonylurea receptor (SUR) gene are
 RT associated with NIDDM in Caucasians."; Diabetes 45:825-831(1996).
 RL [9]
 RP VARIANT PPHI PHE-1386 DEL, AND VARIANTS GLY-1359; ALA-1368 & ILE-1571.
 RX MEDLINE=97081768; PubMed=8923011;
 RA Nestorowicz A., Wilson B.A., Schoor K.P., Inoue H., Glaser B.,
 RA Landau H., Stanley C.A., Thornton P.S., Clement J.P. IV, Bryan J.,
 RA Aguilar-Bryan L., Permutt M.A.;
 RT "Mutations in the sulfonylurea receptor gene are associated with
 RT familial hyperinsulinism in Ashkenazi Jews."; Hum. Mol. Genet. 5:1813-1822(1996).
 RL [10]
 RP VARIANT PPHI ARG-1477.
 RX MEDLINE=96256476; PubMed=8650576;
 RA Nichols C.G., Shyng S.-L., Nestorowicz A., Glaser B., Clement J.P. IV,
 RA Gonzalez G., Aguilar-Bryan L., Permutt M.A., Bryan J.;
 RT "Adenosine diphosphate as an intracellular regulator of insulin
 RT secretion."; Science 272:1785-1787(1996).
 RL [11]
 RP VARIANTS GLN-274; MET-559; ASN-809; CYS-833 AND ALA-1368.
 RX MEDLINE=98178572; PubMed=9519757;
 RA Ohta Y., Tanizawa Y., Inoue H., Hosaka T., Ueda K., Matsutani A.,
 RA Repunte V.P., Yamada M., Kurachi Y., Bryan J., Aguilar-Bryan L.,
 RA Permutt M.A., Oka Y.;
 RT "Identification and functional analysis of sulfonylurea receptor 1
 RT variants in Japanese patients with NIDDM."; Diabetes 47:476-481(1998).
 RL [12]
 RP VARIANTS PPHI
 RP MEDLINE=98282238; PubMed=9618169;
 RA Nestorowicz A., Glaser B., Wilson B.A., Shyng S.-L., Nichols C.G.,
 RA Stanley C.A., Thornton P.S., Permutt M.A.;
 RT "Genetic heterogeneity in familial hyperinsulinism."; Hum. Mol. Genet. 7:1119-1128(1998).
 RL [13]
 RP VARIANTS PPHI PRO-1351; CYS-1419 AND TRP-1492.
 RP MEDLINE=98443218; PubMed=9769320;
 RX Verkarre V., Fournier J.-C., de Lonlay P., Gross-Morand M.-S.,
 RA Devilliers M., Rahier J., Brunelle F., Robert J.-J., Nihoul-Fekete C.,
 RA Saudubray J.-M., Junien C.;
 RT "Paternal mutation of the sulfonylurea receptor (SUR1) gene and
 RT maternal loss of 11p15 imprinted genes lead to persistent
 RT hyperinsulinism in focal adenomatous hyperplasia."; J. Clin. Invest. 102:1286-1291(1998).
 RL [14]
 RP VARIANT PPHI ASP-186.
 RX MEDLINE=99265491; PubMed=10334322;
 RA Otonkoski T., Aemmelae C., Huopio H., Cote G.J., Chapman J.,
 RA Cosgrove K., Ashfield R., Huang E., Komulainen J., Ashcroft F.M.,
 RA Dunne M.J., Kere J., Thomas P.M.;
 RT "A point mutation inactivating the sulfonylurea receptor causes the

QY	DB	Query Match	Best Local Similarity	24.1%, Score 1759; DB 1; Length 1580;	31.8%, Pred. No. 1,7e+106;	Matches 447; Conservative 266; Mismatches 562; Indels 132; Gaps 26;
QY	106	LESCMTFMSLAVAHKKGELSDMEDVWSLSKSHSSDYNCERLERLQEOELNE--VGPD	163			
DB	223	LISKGTYYMMNAFITYAHKK-PIIDRAIGKPLIAMKALNTYORLCEADDAQYKRIQSTQ	281			
QY	164	AASLRVYV-----IFCETRLITSLVCMITOLAGSGPAF--MYKHLEYTOATESNL	215			
DB	282	GA---RAIMQALSHAFR-RILVLSSTFRILADLGFAPLCTGIVDHLGKENDVEYQPKT	337			
QY	216	QY-----SLLVVLGLL--TEIVRMSLALTMALNVRGVRIRGRL	255			
DB	338	QPLGVYFVSSQEFLLANAYVAVLLLEFALLLOFTPLQASLYVAI-----EGGINRGAIQ	391			
QY	256	TPAFKKIILK---NIKESLSGEIINICSDNGOMFEAAAGSLLAGPVAILGIMTYN	311			
DB	392	TYIYKIMHLTSLNSMGMETGQICNLVAIDTQNLMPFEFLCPNLMMAMPQIIVGVILL	451			
QY	312	VIIILGPTGLGAVVITLTPYPMMPFSLRTATYRRCVATDERVQKMEVLYIKFIKMY	371			
DB	452	YVLIVSALIGAIVVILLAPVOYFPAATKLSQAQSTLEYSNRLQOTNEMLGKILKLY	511			
QY	372	AWKAFESQSVOKIREERERILEKAGFQOSITGVAPIVVILSVYTFVSHMTL--GPDLT	429			
DB	512	AMENIFRRVETTRKREKMTSLAPAFATYSISIFMTAIPDAIVALTIFGHSVFKEADPS	571			
QY	430	AAQAFVTVIVFNSMTFALKVTPFSYKSLSEASVAVDRKSLFLMEBVMHKPKKPSPHIK	489			
DB	572	PSVAASLISFLHILVTPFLFSLSSVYRSTKVALVYOKISEFLSSAE--IREDCQAPHP	628			
QY	490	IEKMAVITLAMDSSHSSIQSPKLTFRKMKDKDRAS-RGKKEYROL-----QREHQAVLAE	544			
DB	629	TPQGR-----SKYQVPLRVVNRKRPAEDCKGLTGPQLSLVPSADGDACCQOI	680			
QY	545	OKGHLLDSDERPPSDEERKGIHLGHLRLQTLTSLDLEIGKLVIGICSVSGKTSL	604			
DB	681	MCGYFTWVPDGP-----TLNITFIRIPRGQLTMIVGQVGGCGKSSL	721			
QY	605	ISAILGOMTLLEGSTALS-----CTFNVAAQAMILNLT	638			
DB	722	LIAALGEMKVGAVWSSLPDSEIGEDPSPERETATDIDIRKRGVAAASQKPMPLNLT	781			
QY	639	LKDNLIFGKREYDEERYNSVLNCCCLRPDIALIPSSDTEIGERGANTLGGQQRORISLRA	698			
DB	782	VENITFESPFMKORKKWIIEKQSIQPDIDILPHGQIQIGRGINLSGGQQRORISVARA	841			
QY	699	LYSDRSYTLDDPLSALDAHVGNHIFNSAIRKHLK--SKTVLFVTHQOYLVDCEVIFM	756			
DB	842	LYOHANVAVLEDDPFSALDILHSDHLMQAGILDLRLDKRTVVALVYHKLQYLPHAAWITAM	901			
QY	757	KKGCTIERGTHELMNLNLDYATIFNNLLIGETPPE--INSKTESQSKSDODKGP	813			
DB	902	KGQTIORBSTLDDFORSECOLFEHMKTLTNRRDQOLEKTYTERKATETEPQGLSRAMSR	961			
QY	814	TGSI-----KKRKAYPEBGOLVOLEEKQSGVPSVYGVYIOAAGPPLAFVIALEFNL	868			
DB	962	DGILDDEEBEEBAASEEDDNLSSMLHQRAEIPRACAKYLSAGIILLSTLVSQLLK	1021			
QY	869	NGSTAFSTWMLSTYTKQSGGNTYTRGNETSVSDSMKPNPHQYASTAYALSMAVMILL	928			

Accession	Protein Name	Length	Source
Dd 1022	HMVLVAID-VMLAKTQSDALTLTPAAR	1076	NCISIOECTIDQVYANVFYVLSGLIVL
Qy 929	KAIRVVVVKQKTLRASSSHDELFRRLIRLSPMKPEFDTPTGRILLNRFESKDMEVDVRLPF	988	
Dd 1077	CLVTSVYVEMWGLAKYAKRLHSLNRIILAMRRFETTPGLSILNRFSSDCNTIDQHIPS	1136	
Qy 989	QAEFNIQNVILVFCVGMIAGVFPFWLVAQVPLVILFVSILHIVSVLLIRELKRDLNITQS	1048	
Dd 1137	TLECLSRSTELLCVSLANAVISYVTFEVALPLPAIVCYFIQKFPVVASRDIQQLDDTQQL	1196	
Qy 1049	PFLSHIITSIOGLATPIAHYKNGQSEFLHRYOGLDDNQAPFELFTCAMWLAIVRLDLS	1106	
Dd 1197	PLLSHFAETVGLTTTRFKEYEARQOQLLETPTOSNNIASLFLVANMWELEVRMEYIAC	1256	
Qy 1107	IALLTTTGLMVLVLMHGQIPPAVYAGLAISYAVALGTGLFOFYRLASETEPARFTSVIRIMHY	1166	
Dd 1257	VLLIAAVYISINSILHRELSDALVIGLGYTALMVSIVYLMWYRNADMLQIGAKYRIHGL	1316	
Qy 1167	IKTUSLEAPARIKKNAPS--PDWPEGEVTFENAEKMYRENPLVLKRVSFYTIKPKRKI	1223	
Dd 1317	LKT--EAESEYEGGLAPSLIRPKNMPDQKIQIQLNLSVRYDSSILKVLKHNVALISPGKI	1373	
Qy 1224	GIVGTSGSGKSLGALPRLVELSGCCIKIDGVRISDGLADLRKSLIIQEPVLESGT	1283	
Dd 1374	GICGVTGSGKSSFSFLAFRRWVDTFEGHIIIDGIDIAKLPLRLTKRLSLIILDDPVLFSGT	1433	
Qy 1284	VRSNLDPNQGTEDQIIMPALERTHMKECITAOPLKLESEVMENQGNFSGEROLLICARA	1343	
Dd 1434	IRFNLDPKRSKSDSLWMLALELTAQLKLVKALLPGSLDIITIEGGENFSGOROLFLCLARA	1493	
Qy 1344	LIRHCKILILDEATAAMDTEIDLIOETIREAFADCTMLTIAHRLHLYLGSDRIMVLAQ	1403	
Dd 1494	FVRKTSIFIMEDEATASIDMATEINILQKVVMTAFADRTVVTIAHRYHTLSADLVYLRKG	1553	
Qy 1404	QVDEPDPFSVLLSNDSSRFYAMFAAE 1430		
Dd 1554	ALLEFDKPEKLLSRKDSVF-ASFVRAAD 1579		

RESULT 14

YHDS_YEAST	STANDARD	PRT: 1592 AA.
AC P38735;		
DT 01-FEB-1995 (Rel. 31, Created)		
DT 01-FEB-1995 (Rel. 31, Last sequence update)		
DT 30-MAY-2000 (Rel. 39, Last annotation update)		
DE PROBABLE ATP-DEPENDENT PERMEASE YHL035C.		
GN YHL035C.		
OS Saccharomyces cerevisiae (Baker's yeast).		
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
OX NCBI_TaxID=4932;		
[1]		
RP SEQUENCE FROM N.A.		
RP STRAIN=S288C / AB972;		
RX MEDLINE=94378003; Pubmed=8091229;		
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,		
RA Du Z., Favellio A., Fulton L., Gattung S., Geisel C., Kristen J.,		
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,		
RA Latreille P., Louis E.J., MacI C., Mardis E., Menezes S., Mouser L.,		
RA Nham M., Rifkin L., Riles L., St Peter H., Trevasaks E., Vaughan K.,		
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,		
RA Vandin M.;		
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome		
RT VIII.";		
RL Science 265:2077-2082(1994).		
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).		
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY		
CC (ABC TRANSPORTERS). MRP SUBFAMILY.		
CC -----		
CC This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC the European Bioinformatics Institute. There are no restrictions on its		

```

CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL, U11583, AAB65047.1; -.
DR      PIR, S48933; S48933.
DR      HSSP; P13569; INBD.
DR      SGD; S0001027; YHL035C.
DR      InterPro; IPR001140; -.
DR      InterPro; IPR001617; -.
DR      Pfam; PF00664; ABC_membrane; 2.
DR      Pfam; PF00005; ABC_tran; 2.
DR      PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW      Hypothetical protein; ATP-binding; Transmembrane; Glycoprotein;
KW      Transport.
FT      DOMAIN 1 33 EXTRACELLULAR (BY SIMILARITY).
FT      TRANSMEM 34 54 CYTOPLASMIC (BY SIMILARITY).
FT      DOMAIN 55 74 CYTOPLASMIC (BY SIMILARITY).
FT      TRANSMEM 75 95 2 (BY SIMILARITY).
FT      TRANSMEM 96 100 EXTRACELLULAR (BY SIMILARITY).
FT      TRANSMEM 101 121 3 (BY SIMILARITY).
FT      DOMAIN 122 131 CYTOPLASMIC (BY SIMILARITY).
FT      TRANSMEM 132 152 4 (BY SIMILARITY).
FT      DOMAIN 153 170 EXTRACELLULAR (BY SIMILARITY).
FT      TRANSMEM 171 191 5 (BY SIMILARITY).
FT      DOMAIN 192 329 CYTOPLASMIC (BY SIMILARITY).
FT      TRANSMEM 330 350 6 (BY SIMILARITY).
FT      TRANSMEM 351 379 EXTRACELLULAR (BY SIMILARITY).
FT      DOMAIN 380 400 7 (BY SIMILARITY).
FT      TRANSMEM 401 465 CYTOPLASMIC (BY SIMILARITY).
FT      DOMAIN 466 486 8 (BY SIMILARITY).
FT      TRANSMEM 487 489 EXTRACELLULAR (BY SIMILARITY).
FT      TRANSMEM 490 510 9 (BY SIMILARITY).
FT      DOMAIN 511 572 CYTOPLASMIC (BY SIMILARITY).
FT      TRANSMEM 573 593 10 (BY SIMILARITY).
FT      DOMAIN 594 614 EXTRACELLULAR (BY SIMILARITY).
FT      TRANSMEM 615 635 11 (BY SIMILARITY).
FT      DOMAIN 636 689 CYTOPLASMIC (BY SIMILARITY).
FT      TRANSMEM 690 1010 12 (BY SIMILARITY).
FT      DOMAIN 1011 1051 EXTRACELLULAR (BY SIMILARITY).
FT      TRANSMEM 1052 1072 13 (BY SIMILARITY).
FT      DOMAIN 1073 1115 CYTOPLASMIC (BY SIMILARITY).
FT      TRANSMEM 1116 1136 14 (BY SIMILARITY).
FT      DOMAIN 1137 1137 EXTRACELLULAR (BY SIMILARITY).
FT      TRANSMEM 1138 1158 15 (BY SIMILARITY).
FT      DOMAIN 1159 1229 CYTOPLASMIC (BY SIMILARITY).
FT      TRANSMEM 1230 1250 16 (BY SIMILARITY).
FT      DOMAIN 1251 1252 EXTRACELLULAR (BY SIMILARITY).
FT      TRANSMEM 1253 1273 17 (BY SIMILARITY).
FT      DOMAIN 1274 1592 CYTOPLASMIC (BY SIMILARITY).
FT      NP_BIND 702 709 ATP (POTENTIAL).
FT      NP_BIND 1357 1364 ATP (POTENTIAL).
SQ      SEQUENCE 1592 AA; 180924 MW; 055FB039992AC68 CRC64;

Query Match 23.9%; Score 1749; DB 1; Length 1592;
Best Local Similarity 30.3%; Pred. 7.5e-106;
Matches 439; Conservative 288; Mismatches 495; Indels 228; Gaps 33;

OY      86 SALKRIPTTC-----KQHPVDN--ADLFSCMFTFMSLSLARVAHKKGELSM 130
DB      239 SLTGRIKILSPKIDKSSIIIVYRDDHTFSRSREIISSILCITMSWTINTEW-OKNTIKL 297
OY      131 EDVWSLSKRESSDVNCRRLERLMOEELNEVGDDAASLRVYVIFCRPLLISYCLMTIQ 190
DB      238 KDIMWLSMEDYSIFILKGFTR-NKHIN-----NLTALFESEFKYLLIGMLWVLVNS 349
OY      191 LAGFSGAPFMYKHILEY---QATFSNIOYSLIVLIGLTFEIVWSMLALTMALNTFTG 247
DB      350 IVNL-LPTTLKRFELIYDNPNRSSCNMLMLYITIGMFCIRLTALICNSGQGFVSKIC 408
OY      248 VLRGAIITMAFKLIK-LK-----NIKESLDELINICSGDQRFEEA----- 291

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2001, 17:46:47 ; Search time 108.06 Seconds
(without alignments)
1759.414 Million cell updates/sec

Title: US-09-528-031-2
Perfect score: 7308
Sequence: 1 MMDIDGKRYIIPSPCYRSV.....DSRRFYAMRAAEKNYAVKG 1437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_16:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2371.5	32.5	1427	5 062170	062170 caenorhabdi
2	2082	28.5	1515	4 09099	090999 homo sapien
3	2066	28.3	1545	4 09082	090822 homo sapien
4	2060.5	28.2	1543	11 09118	091188 mus musculu
5	2055.5	28.1	1528	11 035379	035379 mus musculu
6	2038	27.9	1502	6 096K09	096K09 canis famli
7	2022.5	27.7	1547	5 0917N0	0917N0 drosophila
8	2022.5	27.7	1591	5 09VK56	09VK56 drosophila
9	2013.5	27.6	1622	10 024635	024635 arabidopsis
10	2013	27.5	1528	5 09N2N3	09N2N3 c.c. elegan
11	2005	27.4	1540	5 094136	094136 caenorhabdi
12	1993	27.3	1623	10 048907	048907 arabidopsis
13	1992	27.3	1623	10 064590	064590 arabidopsis
14	1988.5	27.2	1525	5 020943	020943 caenorhabdi
15	1984.5	27.2	1494	5 020943	020943 caenorhabdi
16	1972	27.0	1623	10 022449	022449 arabidopsis
17	1959	26.8	1368	10 09S9R0	09S9R0 arabidopsis
18	1953.5	26.7	1622	10 048908	048908 arabidopsis
19	1897	26.0	1573	5 093552	093552 caenorhabdi

20	1882	25.8	1327	5 09VLN6	09VLN6 drosophila
21	1882	25.8	1539	10 09LZJ5	09LZJ5 arabidopsis
22	1865	25.5	1459	4 090Q97	090Q97 homo sapien
23	1858.5	25.4	1456	4 090QAO	090QAO homo sapien
24	1857	25.4	2580	5 09VJ21	09VJ21 drosophila
25	1842	25.2	1465	3 09P5N0	09P5N0 schizosacch
26	1830.5	25.0	1514	10 09LKE4	09LKE4 arabidopsis
27	1828	25.0	1514	10 065619	065619 arabidopsis
28	1819.5	24.9	1516	10 024525	024525 arabidopsis
29	1818	24.9	1545	11 09JU67	09JU67 ratius norv
30	1815.5	24.8	1306	10 09LU34	09LU34 arabidopsis
31	1811.5	24.8	1348	5 09VDR3	09VDR3 drosophila
32	1811	24.8	1515	10 024510	024510 arabidopsis
33	1809	24.8	1513	4 09NXY3	09NXY3 homo sapien
34	1787.5	24.5	1307	5 09VGM1	09VGM1 drosophila
35	1784	24.4	1546	11 09L170	09L170 mus musculu
36	1782.5	24.4	1545	11 063563	063563 ratius norv
37	1781.5	24.4	1549	4 060706	060706 homo sapien
38	1772.5	24.3	1490	10 09MLC7	09MLC7 arabidopsis
39	1768.5	24.2	1493	10 09LKE2	09LKE2 arabidopsis
40	1767.5	24.2	1346	5 09VANS	09VANS drosophila
41	1767	24.2	1144	5 09XW61	09XW61 caenorhabdi
42	1763.5	24.1	1549	6 082451	082451 oryctolagus
43	1752.5	24.0	1582	11 070532	070532 ratius norv
44	1747.5	23.9	1582	11 09EOT0	09EOT0 ratius norv
45	1727	23.6	1344	5 09V572	09V572 drosophila

ALIGNMENTS

RESULT 1
ID 062170 PRELIMINARY; PRT; 1427 AA.
AC 062170; 062241; 09U095; 07, Created)
DT 01-AUG-1998 (TREMBLrel. 13, last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE HYPOTHEtical 158.7 KDA PROTEIN F14F4.3 IN CHROMOSOME X.
GN F14F4.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodrilinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wallis J.M.;
RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Percy C.M.;
RL Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.
CC -I- ALTERNATIVE PRODUCTS: TWO FORMS; ISOFORM F14F4.3A AND ISOFORM F14F4.3B (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC EMBL; AL021446; CAB54225.1; -.
DR EMBL; 283227; CAB54225.1; JOINED.
DR EMBL; 283227; CAB54251.1; -.
DR EMBL; AL021446; CAB54251.1; JOINED.
DR EMBL; AL021446; CAB54226.1; -.
DR EMBL; 283227; CAB54226.1; JOINED.
DR EMBL; 283227; CAB54250.1; -.
DR EMBL; AL021446; CAB54250.1; JOINED.
DR HSSP; P13569; INED.
DR InterPro; IPR001140; -.
DR InterPro; IPR001687; -.
DR InterPro; IPR003439; -.
DR Pfam; PF00664; ABC_trans_2.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_2.
KW Hypothetical protein; Alternative splicing.
FT VARSPLOC 643 669 MISSING (IN ISOFORM F14F4.3A).
SQ SEQUENCE 1427 AA; 158685 MW; 81DA0FBC90961FAE CRC64;

[illegible]

QY	1046	TQSPFLSTISSSTQIGLATHAVKKGQEFHRYQELLDDNQAFPEFTGCMRWLAVRDLI	1105
QY	1046	TSQSPFLSTISSSTQIGLATHAVKKGQEFHRYQELLDDNQAFPEFTGCMRWLAVRDLI	1105
Db	969	SRSPLDYHVSASLEGITTTHTTFQOQSRFLVELKHLIDCNSGAIEMFQSMRWLAVMIDL	1028
QY	1106	SIALITTTGMIYVMGQIPPAVAGALISYAVOLTGLQFOFVRLASETEARTSVYERINH	1165
Db	1029	VYVWTAIVALLTJYMLIGTVSPADAGMALAFVAMQSGIFQFAVRTOLELEAKMTSEVRSY	1088
QY	1166	YIKTSLSEAPARIKNKAP-SPDPQGEVYTFENAEKRYRENIPLVKKVSTFIKPEKIG	1224
Db	1089	YADNIPEDGEMNTRQGLDIESSWPANGQINFSEVNLRYKRSHPILNDITEIKGGEKVG	1148
QY	1225	IVGTTGSGKSSLSGMAFLRVELLSSGCIKIDYRISDIGLADRSKLSIIPPEVYFSGTV	1284
Db	1149	IIGRTGSGKSSLSLNLFRRLPYVNTGTLYIDGVDIRVGLVKLRGISALADQPSLESGTV	1208
QY	1285	RSNIDPPNQYTEDQIMDALEERTHMKCIADLPKLESEWENGDNFSGYEROLLCIARAL	1344
Db	1209	RFNIDPSELEYSDSIMVEALEKCHKLTVLQSLDKKLELADVSHSGNNFSGYEROLLCIARAL	1268
QY	1345	LRRCKIILIDEPATAMDETETDLLIOETIRBPADCTMTLIAHRLHTVIGSDRIWYLAQO	1404
Db	1269	LMKSRIVYILDEAPATASVDAGDKLIGEVIKTVFADAVIILAHRLDNVRMDRIMHLKNGK	1328
QY	1405	VVEFDTPSVLSLNDSS	1420
Db	1329	LINFTEPQEMFKDDWS	1344
RESULT 2			
Q9U099	09U099	PRELIMINARY;	PRT; 1515 AA.
AC	Q9U099;		
DT	01-MAY-2000 (TEMBLrel, 13, Created)		
DT	01-MAY-2000 (TEMBLrel, 13, Last sequence update)		
DT	01-MAR-2001 (TEMBLrel, 16, last annotation update)		
DE	MULTIDRUG RESISTANCE PROTEIN (FRAGMENT).		
GN	MRP.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	GRANT C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;		
RA	Genet C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;		
RT	"Analysis of the intron-exon organization of the human multidrug-		
RL	resistance protein gene (MRP) and alternative splicing of its mRNA."		
CC	Genomics 45:368-378(1997).		
CC	-1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC		
CC	TRANSPORTERS).		
DR	EMBL; AF022853; AAB83979.1; -.		
DR	EMBL; AF022854; AAB83979.1; JOINED.		
DR	EMBL; AF022855; AAB83979.1; JOINED.		
DR	EMBL; AF022856; AAB83979.1; JOINED.		
DR	EMBL; AF022857; AAB83979.1; JOINED.		
DR	EMBL; AF022858; AAB83979.1; JOINED.		
DR	EMBL; AF022859; AAB83979.1; JOINED.		
DR	EMBL; AF022860; AAB83979.1; JOINED.		
DR	EMBL; AF022861; AAB83979.1; JOINED.		
DR	EMBL; AF022862; AAB83979.1; JOINED.		
DR	EMBL; AF022863; AAB83979.1; JOINED.		
DR	EMBL; AF022864; AAB83979.1; JOINED.		
DR	EMBL; AF022865; AAB83979.1; JOINED.		
DR	EMBL; AF022866; AAB83979.1; JOINED.		
DR	EMBL; AF022867; AAB83979.1; JOINED.		
DR	EMBL; AF022868; AAB83979.1; JOINED.		
DR	EMBL; AF022869; AAB83979.1; JOINED.		
DR	EMBL; AF022870; AAB83979.1; JOINED.		
DR	EMBL; AF022871; AAB83979.1; JOINED.		
DR	EMBL; AF022872; AAB83979.1; JOINED.		
DR	EMBL; AF022873; AAB83979.1; JOINED.		
DR	EMBL; AF022874; AAB83979.1; JOINED.		
DR	EMBL; AF022875; AAB83979.1; JOINED.		
DR	EMBL; AF022876; AAB83979.1; JOINED.		
DR	EMBL; AF022877; AAB83979.1; JOINED.		
DR	EMBL; AF022878; AAB83979.1; JOINED.		
DR	EMBL; AF022879; AAB83979.1; JOINED.		
DR	EMBL; AF022880; AAB83979.1; JOINED.		
DR	EMBL; AF022881; AAB83979.1; JOINED.		
DR	EMBL; AF022882; AAB83979.1; JOINED.		
DR	EMBL; AF022883; AAB83979.1; JOINED.		
DR	EMBL; AF022884; AAB83979.1; JOINED.		
DR	EMBL; AF022885; AAB83979.1; JOINED.		
DR	EMBL; AF022886; AAB83979.1; JOINED.		
DR	EMBL; AF022887; AAB83979.1; JOINED.		
DR	EMBL; AF022888; AAB83979.1; JOINED.		
DR	EMBL; AF022889; AAB83979.1; JOINED.		
DR	EMBL; AF022890; AAB83979.1; JOINED.		
DR	EMBL; AF022891; AAB83979.1; JOINED.		
DR	EMBL; AF022892; AAB83979.1; JOINED.		
DR	EMBL; AF022893; AAB83979.1; JOINED.		
DR	EMBL; AF022894; AAB83979.1; JOINED.		
DR	EMBL; AF022895; AAB83979.1; JOINED.		
DR	EMBL; AF022896; AAB83979.1; JOINED.		
DR	EMBL; AF022897; AAB83979.1; JOINED.		
DR	EMBL; AF022898; AAB83979.1; JOINED.		
DR	EMBL; AF022899; AAB83979.1; JOINED.		
DR	EMBL; AF022900; AAB83979.1; JOINED.		
DR	EMBL; AF022901; AAB83979.1; JOINED.		
DR	EMBL; AF022902; AAB83979.1; JOINED.		
DR	EMBL; AF022903; AAB83979.1; JOINED.		
DR	EMBL; AF0		

DR	EMBL:	AJ132287;	CAB45309.1;	JOINED.
DR	EMBL:	AJ245625;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132288;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132289;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132290;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132291;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132292;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132293;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132294;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132295;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132296;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132297;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132298;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132299;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132300;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132301;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132302;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132303;	CAB45309.1;	JOINED.
DR	EMBL:	AJ245626;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132304;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132305;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132306;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132307;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132308;	CAB45309.1;	JOINED.
DR	EMBL:	AJ245627;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132309;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132310;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132311;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132312;	CAB45309.1;	JOINED.
DR	EMBL:	AJ132313;	CAB45309.1;	JOINED.
DR	HSSP:	P13569;	INBD.	
DR	InterPro:	IPR001140;	-;	
DR	InterPro:	IPR001687;	-;	
DR	InterPro:	IPR003439;	-;	
DR	InterPro:	IPR003593;	-;	
DR	Pfam:	PF00005;	ABC_tran; 2.	
DR	Pfam:	PF00664;	ABC_membrane; 2.	
DR	ProSite:	PS00211;	ABC_TRANSPORTER;	UNKNOWN_1.
DR	SMART:	SM00382;	AAA_1;	
DR	SEQUENCE	1545 AA;	174148 MW;	14036686B61628FC CRC64;

[illegible]

D5	551	Q1PPLVLSVYTESVYLVLDVSNMILDAQKAFSTLFLNLRPPLSMWMTSSMLQASVST	610
QY	465	DREKSLFLMEEVHMTKNKPASH-----IKIDMKATLAMPDSSSSSIONSPKLLPKKKD	519
D5	611	ERLEKYLGSGDDL-----DTSAIRHDCNFDKAMQFSEASPTW-----	646
QY	520	KRASRCKKEKVRQLORTBHOAVLADQKGLHLLDSDERSPEEBGKHLHGLRLQRTLH	579
D5	647	-----EHD5-----EATVR	655
QY	580	SIDLEIOEGLVIGCGSVGSGKTSLSIALLGOMTLLEGSIAISGTFAYVAQOAMITLAL	639
D5	656	DVNMDIAQGLAVITPVPGSGKSSLSIAMLGEMENHVHITIKGTAVAPQOSWIOGCTI	715
QY	640	RDNITLFEKYEDEBNYSVYNSOCLRPDLALPESDULTEGERGANTSGGQORISARAL	659
D5	716	KDNILFTEFENEKRYOQVLEACALPDLPEMLRGGLAETGEGKILSGGQKORISARAT	775
QY	700	YSDMSIYIILDDPLSALDAHGNHIFNSALRKH--LKSRTVLFVTHOLOYLVDCEVIFMK	757
D5	776	YQNDIYTLDDPLSADVAHAGKHIFNKVYLPNGLLKGTRLLVTHSMHFLPQVDEIYVLG	835
QY	758	EGCITGERTHEEMLNUNGDAITFNNMLLGETRPVE-----	793
D5	836	NGTIVYEGSVSALLAKGGEPAKMLKFLNHQRPBEEATVHDSGEEDDDYLLISVBEIP	895
QY	794	-----INSKETS-----GSQKSDCKGPKTGST---KKKAYKPREGOLVQ	832
D5	896	EDASAIMRRENSENRRTLSSRSNORHLKSJRNLSIKTRNNVSLKEDDELQVQ--KLIK	953
QY	833	LEEKGGSVPMYGVYVIOAGGPFLEVLNMLFMUNGSTAF--STWMLSYWIKGSGN	890
D5	954	KEFLETKVFASFSTYLETLOAIGLFSIFPITILA-FVNN--SAFISNMLMSM-----	1003
QY	891	TYTVRGNETSVD5MKDNPMPQYASIVALSAMVILIKALRGV--VFVK-----GLRA	943
D5	1004	-----TDSQKIFN--STDYPASQROMRGVYGALGAGIEVFVFIAPHM5AFGEVHA	1052
QY	944	SSRLHDELFRILRSPKKEFDTPPTGRILNRSKMDDEVNVLPRQALMFQNTIYVFC	1003
D5	1053	SNMHLKOLLNNIRLAPMRPFDDTPTRIRYNRPFRAGDSTYDITLPLSLSWITCFLGIST	1112
QY	1004	VGMVAGVPMFLVAVGPFVLTFEVLHIVSRVYLIRELRLDNTQSGPFSHTSSIOGLAT	1063
D5	1113	LVMICMATEPVFTIIVIPDLGITIYVSOMFVYSRSQRLRLDSTYTRSPITSHRSEVSGLPV	1172
QY	1064	IHAVNGOEFLHRYOELLDNQAPFLFTCANRWLAVRLDLISALITTTGMLTVLMHQ	1123
D5	1173	IRAREHQORFLKHNHVAIDPNOQCFVSWITSNMWLAIRLELVGNLTVEF5ALMVIYDRT	1232
QY	1124	IPPAVAGLAI5AVOLTGLEFOFTVRLASETEARFTSVENINHYITLSLEAPAKKNAP	1183
D5	1233	LSGDTVGVL5NALNITQTLNMLVRSMTSELEINIVAVERTIYTK-VENEAP--WYTKRP	1290
QY	1184	SPDPPOBESEVFEHAERBYRENTPLVKVYSFTIKRKEKIGVGTGSGKSLGALFRL	1243
D5	1291	PPDMP5SGKIQFNFYQVRYRPELDVLVRIGTCDISMEKIGVAGRTGAGKSSLNCFLRI	1350
QY	1244	VELSGGCKIIDIVRISIDIGIADLRKSKISIIPOEPVLFSGTV5NLDPNQVETQOIMDAL	1303
D5	1331	LEAAGGQIILIDGVDIA5IGIHDLRKLTITIPQRIIFSSSL5LMNLDPPNNYSDEIKMAL	1410
QY	1304	ERTHMKECIADLPKLESEVWENGDNF5VGEROLCICARALLRHCKLILDEAPRAMDTE	1363
D5	1411	ELAHKISFVASLQIGLSHEGTEAGNLSIGORQLDGLGALLRKSKIILVLEAPAVADLE	1470
QY	1364	TDLLLOETIRFAPADCTMLTIAHRLHTVLGSDRIKVLADGQVAVERTDTSVLLSND5--RF	1422
D5	1471	TDNLIQOTIQQEFAHCHVITIAHRLHTIMDSKVWVLNDGKILIEGSGPEELLQIPGPPTF	1530
QY	1423	YAMFAAENKVAVK	1436
D5	1531	MAKEGJENVN5TK	1544

OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Conrad S., Viertelhaus A., Orzechowski A., Hoogstraate J., Gjellian K.,
 Schrenk D., Kauffmann S.;
 RT "Sequencing and tissue distribution of the canine MRP2 gene compared
 with MRP1 and MDR1."
 RL Toxicology 0:0-0(0).
 DR EMBL: Y18220; CAC17701.1;
 SO SEQUENCE 1502 AA; 169265 MW; A01B26D946C0EBBC CRC64;

Query Match 27.9%; Score 2038; DB 6; Length 1502;
 Best Local Similarity 34.7%; Pred. No. 7.3e-130;
 Matches 488; Conservative 239; Mismatches 475; Indels 204; Gaps 25;

OY 99 HGVNDAAGFSCCTFWSLSLAVAKKGLSMEDVMSLSKHESSDVNCKRLERLMOEL- 157
 DB 193 NSTFASLSTSTFWSDIYMKYKQ-PLTLEDVWDDEQITTKALVSKFEKYVEELQ 251
 OY 158 -----NEVGPD-----A 164
 DB 253 KAKTKLOKQOQKNTQKSGSERLHDLNKNSODILVLEVKKKKSGTTEKFKPSMLV 311
 OY 165 ASLRVNWIECTRLILSLICMTQLAGSGPAPWKHLEYQATE---SNLOYSLL 220
 DB 312 KSLFTEFYI---LTKSFLKLVFDLTFPLNPQ-LTKLLISPANDPMYVMTGYFYSVL 366
 OY 221 LVGLLTFIVSWSLALTMALNTRGVRLRGAILLMAKTKLTKLN- IKESLGEELN 278
 DB 367 ----FFVALIOSLCQSYFQMCFLGVNVRTIMASTYKKAULTSNOARKYTTIGEYV 422
 OY 279 ICSNDGQRMFEAAVAGSLAGPVAALIGMAYNIIIGPTGFSVFIIFYPAMFASR 338
 DB 423 LMSVDAQKMDVTNFIHLSMNVLDALSLIYFLMELGFSILAGVVMILLIPVNDLLAS 482
 OY 339 LTAVRKRCVAATDERVQKQNEVLTYIKFIKMYAMVKAFSOSVQKIREBERILEKAGYF 398
 DB 483 KSRALQVKNMKKDKRLIMNEILSGIKILKYFAWEPSPKNOVHEIRKKELKMLTFGQM 542
 OY 399 OSITGVAAITVVIVASVYFVSVMHLGPD--LTAQAFTVYVFNMTALKYTPSVVS 456
 DB 543 OSVMVFLVLTPLVLSVITFVSYYTLVDSNNVLDAAEKAFSTLFLNLRPLSLMPVVIS 602
 OY 457 LSEASVAVDRFKSLFIEMEVNHIK-NKPASPHIKIEMKANATLAMDSSHSSIQNSPKLTPK 515
 DB 603 LLOASYSRRLKLYLGDDLDTSAIRDSSSDKAVOFSEASFTWD----- 647
 OY 516 MKDKRASGKKEKVRQLOLRTHEQAVLAEQKGLHLLSDSPRSPEEESKNIHLGLRLQ 575
 DB 648 --RDSEA----- 652
 OY 576 RFLHSIDLEIOGKLVIGSGVSGKTSLSIALGOMTLEBSIALSGFPAYAAQOAML 635
 DB 653 -TTRDVNLEIMGLMAVAVGVGSGKSLMSAMLGEMEVHGIITTKTAYVPOOSWIO 711
 OY 636 NATLRDNLIFGKEDERYNVLSNCSCLRPDLALIPSSDLTEIGERGAINLSGQORISL 695
 DB 712 NGTINDNLIFGSELDERRKQOVLLEACALLPDLLEVLPGLAEIGEGINLSGQORISL 771
 OY 696 ARAALYSRDSIYILDDPLSLADAHVGNHIFNSAIRKH-LKSKTVLFVTYHQLOLYVDCVE 753
 DB 772 ARAATYNSDIYVLDPLSLADAHVGNHIFNKYVGPGLLKGKTRILVYTHSHPLPYDEI 831
 OY 754 IMKKECITIERGHEHMLNLNGDIATIFNNL-----LGETPPPE 793
 DB 832 VVVGNGTILKESISYNNLLAKKGLFAILATFKYONGPEEATVNEDESDDDCGIMPSEVE 891
 OY 794 -----INSRKETS-----GSQKKSQDKGPKTGSIK-----KEKAVPEEGQ-L 830

DB 892 EIPEEVAALTMRENSLHRTLSRSSRSRHRQKSLNSLKTNRVNTLKREE-EPVAGQL 950
 OY 831 VOLEEKGGGVPMWYGVYITQAAAGPLAVLYMALFMLVNGSTASTWMLSTWIKGSGN 890
 DB 951 IKKEFQTKVKEFSIYKLYRAIGWLIPLIIFA-VYINSVAYIGSNLMSAMTNDSKA- 1008
 OY 891 TTVTRGNETSVSDSMKDNHMOGYASIVALSAMVMLILKAINGVYVVKTKLRASSRLHE 950
 DB 1009 -----FNGTNPASQDMRIGY--GVGLAGGVFLM--ANLSAHSHTASNLHMQ 1058
 OY 951 LFRRLIRSPKFEEDTPTTGRIILNRESKDMEVDVRLPFOEMETQVNLVFCVGMIAV 1010
 DB 1059 LNNLILQAMSFEDPTPTGRIYNNRFAAGDLSYDDLPQSLRSLILCEFLIVSLVMICRA 1118
 OY 1011 FFWFLVAGPLVILSVLHIVSRVLIRELRDNTQSPFLSHITSSIOGLATIAVYKNG 1070
 DB 1119 TFWFIVIIPLSLIYVSIOIFVAVNSROLKRLDSVYRSYSHFSEVSGLSYTRAFHQ 1178
 OY 1071 OEFRLHYOELDNDQAPFLFCAMRLAVRLDISALITTTGLMIVLMHQIIPAYAG 1130
 DB 1179 QRLKHEVGDIDNQCVRSMVSNRMLAVRLIGNLVLFSSLMVLYKATLSGDIVG 1238
 OY 1131 LAISYAVOLTGLEFQFTVRASETEARFTSVIRINHYIKTSLSEAPARINKAPSPDMPQE 1190
 DB 1239 FVLSNALNTQTOLNMLVRYMTSELETVIVAVERINEK-VENAP-WYTDKRPPEWPKR 1296
 OY 1191 GEVTFENAMRYRENPLVLYKAVSFYIKREKIGIVRTGSGKSSGLMALFRLVELSGC 1250
 DB 1297 GEIRFNNYQVVRPELDLVLKRGITCDIRSEMEKIGYVGRGACSGSLTNLDFRILEMAQGO 1356
 OY 1251 IKRDGVRIIDIGLADLRSLSTIIPQEVPLFSGVRSNLDLPFNOYTEDQIMDLERTHMKR 1310
 DB 1357 IITDGVDISIGLHDKRLKTLIIPDPIIFSGSLRNKLPFNHYSGELWKALELAHKT 1416
 OY 1311 CIAOLPLKLESEVMENGDNFSVGERQLCIARALRHCKILLIDETAAMPDETLLIOE 1370
 DB 1417 FVAGIQGLSHFEVAEAGDNLSIGORQLCLLARLLKSKTILIMDEATAVADDETDLIOM 1476
 OY 1371 TIREAFDCTMLTIARHTVAGSDR 1396
 DB 1477 TIQERSHCTTITIAHRLHTIMDSK 1502

RESULT 7
 ID 0917N0 PRELIMINARY; PRT; 1547 AA.
 AC 0917N0;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
 DE CG6214 PROTEIN.
 GN CG6214.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins K.A., Galie R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Agbayani A., An H.-J., Andrews Pflankoch C., Baldwin D.,
 Balley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Diez S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischman W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jammal B.E., Kodali C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Mikhina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Splet E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Stvikas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA The genome sequence of Drosophila melanogaster.,
RA Science 287:2185-2195(2000).
CC - SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
TRANSPORTERS).
DR EMBL: ABO03637; AAC22430.1; -.
DR InterPro: IPR000719; -.
DR InterPro: IPR001140; -.
DR InterPro: IPR001687; -.
DR InterPro: IPR003439; -.
DR Pfam: PF00005; ABC_tran; 2.
DR Pfam: PF00664; ABC_tran; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR ATP-Binding; Transport.
SQ SEQUENCE 1547 AA: 173190 MW: 492894505B80681 CRC64;

Query Match 27.78; Score 2022.5; DB 5; Length 1547;
Best Local Similarity 33.08; Pred. No. 8.8e-129;
Matches 497; Conservative 256; Mismatches 455; Indels 299; Gaps 34;

80 KYHGLSALKPRTCTQHVPDAGLFCSTFSLSLARVAHKKGELSMEDWSSSKH 139
172 KYQNG-----ENRPELSASFLSRTTYQFPDKMALKGR-RNPLEKDKMDLRPQ 219
140 ESSDVCNCRRLERLMOEELN-----VGPD-----ASLRR 169
220 DSCSEVPPIFAHNNQNVKRYKKARKEPKAOPNSGNVTPEFHGKNGKRMASIMP 279
170 VVMFECFTRILTSIVCLMQLGAGSPAPVKKLLLEYQTQTESNLQYSLLVGLILTE 229
280 PIYSEGVFPEGALMKLEFDLTLPAPQ-VLSLISIVEQDEPEPK-----GILYAV 333
230 IVRSMSLALPRTCTQHVPDAGLFCSTFSLSLARVAHKKGELSMEDWSSSKH 281
334 LFLYLAAGQFLLIQGYHFMIVGLRIRTLALINAYRKALISNSTKESVGEIVNMA 393
282 NDGGRMFEEAAGVSLAGPVVAILGMIVNYILGPTGFGSAVFLEFPAM-MFASRLT 340
394 VDAQREFELFTYLLMIMSAFLQIGLALYFLMOQLGSPVLAGLAVIILIPVNGVIASRIK 453
341 AYFRKKAATDENVOKMEVLTATIKFKMAYWAKASOSQKIREERRILLEKAGYFOS 400
454 TYQIRQ-MKYDERVKLMEVLSGIVKLTAMPESPFKOVLDIRKEIATLRSTAYLNA 512
401 IT-----VGADIVVIA----- 413
513 GTSLMSCAPFLVSLVTFATYVLTSEANQLSEKIVIVISIALFDLMKPLRTLLPMASVDA 572

QY 414 --SVYFSVHMTLGF--LTAQAFTYVYVNSMTFALKYTPPSVSLASPAVDRFKS 469
DQ 573 EVSIVTFATYVLLDENNVLDATATFVSLISFNILRPELTMILITNLVQVSVNINK 632
QY 470 LFLMEEV--HMKNKPAKSPKIKEMKATLMDSSHSIONSPEKLPKMKKRASRGK 527
DQ 633 FLNSELDPVSLHDSKPH-PMSENGEFSW----- 663
QY 528 EKVROLRTEHQAVLAQKQHLHLDSDRSPSEEEKHILHGLRLQRTLSIDLEIQE 587
DQ 664 -----GDEI-----TLRNINIEVK 678
QY 588 GKLVIGCGSVSGKTSLSIALQMTLBSISALSGFAVVAQQAAMTLNIDNIFGK 647
DQ 679 GSTVALGTGSGSVQAFLGEMELAGVNTVGLAVPQDAWQNTKRNILLEG 738
QY 648 EYDERYNSVNSCCLPDLALPSSDLTEIGERGALSGGQORISLARALYSDRSIYI 707
DQ 739 TYDKRYNKYIDCALADADILSAGDLTEIGERGINLSSGQORISLARALYSDADLYL 798
QY 708 LDDPLSALDAHVNHNISAL--RKHLKSTVLEVTHQQLYVDCDEVIFMKREGCITERG 765
DQ 799 LDDPLSALDAHVGKHIPEEYIGPKILARKSRVLTGVTFLPQVDIYVYIKMEIISG 858
QY 766 THEELMINDYAL-IRNNILIGETPPEVINSK-----ETGSGQK----- 806
DQ 859 TFDQLVKNKGFADFIQLHLEGNEEELNQIKRQISSADVPPELLGYEKAIKARTE 918
QY 807 -----SODKPKTGSIKKERKAVKPEGVQVLEK 837
DQ 919 SLDSISVTSNDSLSMGGGSLRRRTKQSDSHSVASAASLKKQEV--BGLKLETRKSQ 975
QY 838 QGSVPMSVGYVYIQAAGCPALFVLMALMNGSTAF--STWWSIYWKQSGNTVT 894
DQ 976 TGGEVFAVYKHYKISVG----IFLSVATLVLFVFOAFQJSGNSMLTQW----- 1020
QY 895 RGNETSVDSSKDNPHQYASIAL-----SNAVMILKAIKGVVVKTLR 942
DQ 1021 -ANDQNVAN--DTGLRDMVLGYGAFGGVATNFSSLAISL-----GCLK 1064
QY 943 ASSRLHDELFRILRSPMKPFDTPPTGRILNRFSDKDEVDVRLPFOAEMIQNVILVF 1002
DQ 1065 CSQILHQTLLYNNLRWPELEFDTTPGLRIYKRFESKDITDNLVFNIRVYIGAYVLA 1124
QY 1003 CVGMIAGVFWFLVAVGPVLVLSVLHIVSVLI--RELRKLDNITQSPFLSHITSIQ 1059
DQ 1125 TIVVISLSTPIFLAVIVPIALY--YFAQRFYATRSQMLRELSVRSPIYSHSETVT 1181
QY 1060 GLATTHAVNKGQELHRYQELLDONQAPFLFTCAMWLANRDLISALITTTGLMIVL 1119
DQ 1182 GASPIRAYNGVDRIEESDAKVDKNQCKYPSVIANRWLAIRLEMVGNLILFASLFAVL 1241
QY 1120 MHGQIPPAYAGLAIYAVQLTGLFQFTVRLASFTSEARFTSVIERINHYIKTSLPARIK 1179
DQ 1242 -GGQTNGLVGLSVYALQVYQTLNWLVRMSSDIETNIVSVERIKKEGET-KQAPAPLE 1299
QY 1180 NKADSP-DWQEGEVTEENAMRYRENPLVLKRVFTIKPKKEIGYVGTSGSKSLGM 1238
DQ 1300 QDKRKKMPQEGEVFEONFQYVRREGIDLVRGVSINQGEKGVYIGRGAKSSITL 1359
QY 1239 ALFRVLVELSGGCIKIDGVRISDGLADLRSLTIIPQPVLFSGTGSVNSNDPEQYEDQ 1298
DQ 1360 ALFRILAEAGRISIDGVDIASMGLHMRSLTITPDPVLFSSGLRNIDPEFIKTDE 1419
QY 1299 IWDALERTHMKECIAOLPLKLESEVMGDNFVSEROLICIAALRHCKIILIDETA 1358
DQ 1420 IWKALDELHLSKFSYKSLAAGNHIAEGEMLSGVQVQLARALLRKRVVLIDEATA 1479
QY 1359 AMDTETDLLIQETIREAFADCTMTLTAHRLHTVLSDRIMVLAQGVVEDPDSVLISND 1418
DQ 1480 AVDETETDLLIQ-----FKECTVLTIAHRLNTIIDSQVIVLDGQIIEFASPELIDNP 1533

0y	1419	SRFRYAM	1425	
		11111		
Db	1534	KSATFSM	1540	
RESULT	8			
09VK56				
AC	09VK56	PRELIMINARY;	PRT;	1591 AA.
DT	01-MAR-2000	(TREMBLrel, 13, Created)		
DT	01-MAR-2001	(TREMBLrel, 16, Last sequence update)		
DT	01-MAR-2001	(TREMBLrel, 16, Last annotation update)		
DE	CG6214	PROTEIN.		
GN	CG6214			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_Taxid=7227;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BERKELEY.			
RX	MEDLINE-20196006; PubMed-10731132;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champagne M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abell J.F., Asgari A., An H.-J., Dettl R., Nelson C.R., Baldwin D.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Duran-Rocha S., Dunkov B.C., Dunn P.,			
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jatall M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lascko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Mishina N.V., Moberly C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclty J.M.,			
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,			
RA	Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,			
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,			
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,			
RA	Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,			
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;			
RL	"The genome sequence of Drosophila melanogaster.";			
RL	Science 287:2195(2000).			
CC	-1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC			
CC	TRANSPORTERS).			
CC	EMBL; AE003637; AAF53223.2; -;			
DR	HSSP; P13569; 1NDB.			
DR	FLYbase; FBgn0032456; CG6214.			
DR	InterPro; IPR000719; -;			
DR	InterPro; IPR001140; -;			
DR	InterPro; IPR001687; -;			
DR	InterPro; IPR003439; -;			
DR	Pfam; PF00005; ABC_tran; 4.			
DR	Pfam; PF00664; ABC_membrane; 3.			
DR	PROSITE; PS00211; ABC_TRANSPOFTR; 2.			

[illegible]

QY 895 RGNETS DSKNDPNHMYASYAL-----SMAVMILKAIKGVVFKGTLR 942
DB 1065 -ANDQYAN--DTGLRDMYLGYGAFEGQVATNFSSALISL-----GCLK 1108
QY 943 ASSRLHDELFRRLIRSPKPFEDTPTGRIIRNSKDMDEVVRLPFOAEFICVILVVF 1002
DB 1109 CSQLLHQTLLLYNRMWMEFEDTTPGRIRYNRSKDIIDTIDNVLPFMRVIGAYVLA 1168
QY 1003 CVMGIAGVFWFLVAVGPLVILSVLHIVSRVLI---RELKRDNTIQSPFLSHITSIQ 1059
DB 1169 TIYVISTSTPIFLAVIPIAFVLY---YFAQRFYVATSRQIMRLSESRSPIYSHSETVT 1225
QY 1060 GLAIIHYNNGQELHAYQELLDONAPFLFCAMRWLAIVRLDLSIALITTTGMLIVL 1119
DB 1226 GASTIRAYNNGDRFIEESDAKVDKNQCKYPSYANRMALRLMVGNLITLPAASLPAVL 1285
QY 1120 MHGQIPAYAGIAISYAVOLTGLFQFTVRLASETAEFTSVERINHVIKTLSLEAPARIK 1179
DB 1286 -GGGTNPGVGLSVSYALQYTLNMLVRSMDLETNIVSVERIKEVGET-KQAPMELE 1343
QY 1180 NKASP-DWPGEGEVTEENAEEMRYRENLPVLKVSFTIKPEKIGIVGRTSGKSSILGM 1238
DB 1344 QDKKPKNWPQEGVEFQNFQVRYREGIDLVRGVSFNIGGKGVIGRTGAKSSITL 1403
QY 1239 ALPLVELSGGCIKIDGVRISDGLADLRKSLTIPOEPVLFSGTVSNIDPFNOYTEDQ 1298
DB 1404 ALFPIIAAGRISIDIVDIAWSGLHMRSLRTIIPQDPVILSGSLRNLDPEFKITDDE 1463
QY 1299 IMDLEERTHMEKCIQAQPLKLESEVMENGDNFSGVEROLLCTARALLRHCKIILDEXTA 1358
DB 1464 IWRKLELSHLSKSPKSLAAGINHEIAGENGELSVGQROLVCLARALLRKTRVILVDEXTA 1523
QY 1359 AMDTETDLLIQETIREAFADCTMTLTAHRLHTVLGSDPRINVLAAQGVVEFDPVLLSND 1418
DB 1524 AVDETFDLIQ-----FKECTVLTIAHRLNTIIDSQKVIYLDKQILIEFASPELIDNP 1577
QY 1419 SSREYAM 1425
DB 1578 KSAPYSM 1584

RESULT 9
ID 024635 PRELIMINARY; PRT; 1622 AA.
AC 024635;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE MULTIDRUG RESISTANCE-ASSOCIATED PROTEIN HOMOLOG.
GN MRP1.
OS Arabidopsis thaliana (Mouse-ear Cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI-Taxid-3702;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE-97368351; Pubmed-9223346;
RA Lu Y.-P., Li Z.-S., Rea P.A.;
RT "MRP1 gene of Arabidopsis encodes a glutathione S-conjugate pump:
RT isolation and functional definition of a plant ATP-binding cassette
RT transporter gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:8243-8248(1997).
RN [2]
RN SEQUENCE FROM N.A.
RP Lu Y.-P., Li Z.-S., Rea P.A.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RA Lu Y.-P., Li Z.-S., Rea P.A.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
TRANSPORTERS).

DR EMBL: AF008125; AAB71832.1; -
DR EMBL: AF008124; AAB67319.1; -
DR HSSP: P13563; INBD.
DR Mendel: 24381; Arabid.1944;24381.
DR InterPro: IPR001140; -
DR InterPro: IPR001687; -
DR InterPro: IPR003439; -
DR InterPro: IPR003593; -
DR Pfam: PF00005; ABC_tran. 2.
DR Pfam: PF00664; ABC_membrane. 2.
DR PROSITE: PS00211; ABC_TRANSPORTER. 2.
DR SMART: SM00382; AAA. 1.
KW ATP-binding; Hydrolase; Transport.
SQ SEQUENCE 1622 AA; 182015 MW; 6F0287AF651AD5CF CRC64;

Query Match 27.6%; Score 2013.5; DB 10; Length 1622;
Best Local Similarity 34.3%; Pred. No. 3.9e-128;
Matches 476; Conservative 240; Mismatches 501; Indels 171; Gaps 21;

QY 74 EEPKGYHHGLSALKPIRTTCQHQVNDNAGLFSCMTFSWLSLAVAKHKGELSMEDV 133
DB 217 EELPGGE-----NICEPRH---ANLPDSIFFSWLNLMTLGSKR-PLTEKDV 259
QY 134 WSLKHSNDVNCRLRLRMOEELNEVGPDAASLRVYWIICPTRLILSTYCLMITQLAG 193
DB 260 WHLDWTMDKTETLMSFQSWMDKELEKRP-----W-----LIRALNLSLG 300
QY 194 -----FSGPAFWKHLELYQATES---NLQYSLLVGLLLEIVRSW 234
DB 301 RFWMGFWKIGNDSCSOFVGP-LINELLKSMQNLNEPAMIGIYVLSFVGVLGICE-- 357
QY 235 SLATMALNTGTGVRLLGAILTMAFKILKIKTKESL--GELINICSDNGQMFEPA 292
DB 358 --AQYFQVMVNGVRLSALTAFAVRKSLRLTNEGRRKFKQGTNTMTDAESLQIQ 415
QY 293 VGSLLAGPVAIIIGMIVNIIIGPTGLGSAVFLFPAAMFASRLFAVFRRCVATD 352
DB 416 SLHTWMSAPFRITVALVILYQOLGVAISIGLFLVLFPIQTVIISTQKLTGKIGRTD 475
QY 353 ERVQKMEVLTIRIKIKYAWAVKAFSOSVQKIREERRILEKAGYFOSITGVAPIVVI 412
DB 476 KRIGLMNEVLAAIMDTVVCYAMENSFOSKVQTVRDELSWFKKQLLSAFNNFILNSIPVL 535
QY 413 ASVTFVSHMTLGFDLTAQAFVTVVFNSTFPAIKVTPESVKSLSASVAVDFKSLFL 472
DB 536 VTVVSFGVSLGDLPARAFSTLSLFSVLRFLPMIPNITQMVNANVSLNLEVL 595
QY 473 MEYHMIKKNKASP-HIKIEKNATLMDSSHSIIONSFKLTPMKKDKRASRGKKEVR 531
DB 596 TEERVLLPNPIEPGQPAISIRNGYFSWDS----- 625
QY 532 QLORTHQAVLAEQGHLLDSDERSPEEECKHILGHLRLQRTLHSLDLETOEGKLV 591
DB 626 -----KADRP-----TLSNINLDIPLSLV 645
QY 592 GIGSGVSGKTSLSAILIGM-TLEGSIAISGTPAYVAAQAMLNATLNPILFGEYD 650
DB 646 AVVGSTGKTSLSISAMIGELPARSDATVTLRGSAVAYPVQSWIFENAVRNILLEGAFD 705
QY 651 EERNYSVNSCCLMRDLALPSSDLTEIGERGAMLSGGORISLARVLSDRSYTID 710
DB 706 QEKERYVIDVTAIHDDELPLGGDLTEIGERGVNISGGQKQVSMARVYNSNSDCLIDE 765
QY 711 PLSALDAVGNHIFNSAIRKHLKSKTVLFVHOLQYLDCEVIFMKEGCITTEGTHEEL 770
DB 766 PLSALDAVGOVEKCKRKRELQGTTRVLVYNQHLFSQVDKILLVHGTVKEGTYEEL 825
QY 771 MNLNGVATITNNL-----LIGETPPVEINSKKEKTSQSKSODKPGTGSIKKEKA 822
DB 826 CHSGPLFRLMENNGKVEDYSEENGAEVHOTSVKPVENGANNILQKQGITKSKSENS 885
QY 823 VKPEGQVLQLEKGGGGSVPWSVGYTQAGGCLPLFLVIALPMLNANGVAFSTWMLSY 882

```
Db 886 v-----LVKREERETGVVSMKVLERYQNALGAMVVMVLVICYVLTQVFRVSSITWLS 939
Oy 883 WKOGSGNTTVRGNTSVSDSMKDPHMOYASIALSMANVLLIKARGVVFGTJR 942
Db 940 WIDSG---TPKHG-----PLEFYNIVALLSFGVSVTLINSLWLNSSLY 982
Oy 943 ASSRLDELFRRLISSPMKFPDPTPTGRILNRFSSKMDVDVLPQAEFIONVILVFE 1002
Db 983 AKKXHDALGSLRPMVFFOTNPLGRILNRFKAKMGIDIRVANFVNMFKSTIQLLS 1042
Oy 1003 CVMGAGVPMFVLVAVGPLVILFVSLVHSVRLIRELKRDLNITQSPFSLHTTSIQGLA 1062
Db 1043 FYVLIGIVSTLSLMAIMPLLVFVYGALYYQNTSREIKRMDSTTRSPVYAQFCEALNGLS 1102
Oy 1063 THAYNKGEPFLHRQELLDDQAPFELFTCAMRLAVLNLISLITTTGMLVLMG 1122
Db 1103 STRAKVADMAEINGRSMDNNIRFTLVNMANRMGLIRELVGGLMVWTASLAVMONG 1162
Oy 1123 QT--PPAYV---GLAISYAVOLTGLFQFTVRLASEFEARTSVERRINHYIKTSLAPAR 1177
Db 1163 KANQOAVASTMGLISLSTSLTAVRLASLAEINSLNVERVGNITELPS--EAPLV 1221
Oy 1178 IKKAPSPDMPQEGEYTFEMAEKRYRENDPLVKVYFTIKPEKIGIVRTGSGKSLG 1237
Db 1222 IENNRPPGMPSSGSIKFDVYLKRYRELPVYLHGVSFLLSPMDKYGIVRTGAGKSSLL 1281
Oy 1238 MALFRLVELSGGCIKIDYRISDIGLADRSKLSITPQEPVLFSGTVRSNLDPNQYTED 1297
Db 1282 MALFRLVELSGGCIKIDYRISDIGLADRSKLSITPQEPVLFSGTVRSNLDPNQYTED 1341
Oy 1298 QLMDELTPHMEKCTIAOLPLKLESEVMENGDNPSVGEORLLCARALLHCKLLIDEAT 1357
Db 1342 DLMESERHMLDQTKRNPGLDAEYTEAGENFSGORLLSLARLLRSKLVIDEAT 1401
Oy 1358 AAMDTEIDLIOETREAFDCTMLTIAHRLHTVLSGRIMVLAQOVVEFDTPSVLLSN 1417
Db 1402 AAVDRTVDLQIKTIREERKSCMTLIIARIIMTIIDCKRVIVDSKQVDFSSPENLISN 1461
Oy 1418 DSSREYAM 1425
Db 1462 GESSFSKM 1469

RESULT 10
OQ9N2N3 PRELIMINARY; PRT; 1528 AA.
AC OQ9N2N3;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
DE C. ELEANS MUTLIARUG RESISTANCE RELATED PROTEIN 1 (MRP-1) (GB:U66260),
DE CONTAINS SIMILARITY TO PFAM DOMAINS PF00664 (ABC_MMBRANE,
DE E-VALUE=1.3E-54, N=2) AND PF00005 (ABC_TRAN, SCORE=277.3,
DE E-VALUE=6.3E-80, N=2).
GN MRP-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chromadorea; Rhabdilitida; Rhabditoidea;
OC Rhabdilitidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium. ";
RL Science 283:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Favello A.;
RT "The sequence of C. elegans cosmid F57C12.";
```

```
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
CC TRANSPORTERS).
DR EMBL: U41554; A031550.2; -
DR InterPro: IPR001140; -
DR InterPro: IPR001687; -
DR InterPro: IPR003439; -
DR InterPro: IPR003593; -
DR Pfam: PF00005; ABC_tran; 2.
DR Pfam: PF00664; ABC_membrane; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR SMART: SM00382; AAA; 1.
KW ATP-binding; Transport.
SQ SEQUENCE 1528 AA; 171046 MW; 3CFC5D831C0EAD0A5 CRC64;
```

Query Match 27.5%; Score 2013; DB 5; Length 1528;
Best Local Similarity 33.9%; Pred. No. 3,8e-128;
Matches 488; Conservative 239; Mismatches 483; Indels 228; Gaps 28;

```
Oy 96 KIOHPVDNGLPSCMFWLSSLARVAHKGELSMEDVMSLSHSSSDVNCRRLERLMOE 155
Db 204 KNSCEPYTAFSLQQLTFQWFGSLAYIGNKK-SLEKEDLDLNRD-----KAENIIPS 255
Oy 156 ELNEVSPDASLR-----VWIFCRTL-ILSTVCL-MITQLAGFSG 196
Db 256 FIENLIPVEGYRRKIKKNPEAIPKNHSILIPITKTKFTLLAGCYKIMMDLLQFA 315
Oy 197 PAFVAKHLLTYQATESNLQYSLVGLLEIYVSWSLATWALNRYGVNRCALIT 256
Db 316 PE-ILQLISFIEDKQNPWIGVSIALLMFLSLSLQSMILHGFHMFRLGNIRSLVS 374
Oy 257 MAFKTLTKLN--IKRSIGELINICSDGQRFEEAAGVSLAGPVAAILGMITNVII 314
Db 375 AVYTKTLNLSNEARKGKTGTAIVNMSVDIQRIQDTFTIMFWSAPLIDILSLYFLMKL 434
Oy 315 LGPTFLGSAVFLTFYPAAMFASRLTAFFRRKCVATDERVQKMFVLYIRKIKYAMV 374
Db 435 LGVSVLAGFVILLILIPNSFISVKMRNCOMEQMKRDERIKMSSELINGMKYLKYSME 494
Oy 375 KAFSGSVOKIRREERRILKAGYFOSITY--GVADIVVIVASVTFVSHMTLGFDP--L 428
Db 495 KSMKKVLEVREREIRVLLKLSYLNATLISMACAPFLV--AVLTFGLYVLDPENNVYL 551
Oy 429 TAAQATTVTVFRSMTFALKVYTFYSKLSSEASVAVDRKSLFLMEEVHMKRKPSPH 488
Db 552 TPQITVALALFELTFPLAVFAMFSAOVCSASWTRKEFPAAEEM-----SPQT 603
Oy 489 K-----IEKNATLMDSSHSISQNSPKLTPMKKDKRASRCKKEKVRQLQRTED 539
Db 604 STAYGTDSAIKMDGGSFAM-----GSKEE----- 628
Oy 540 AVLAQKGHLLDSDRPSPEEERKHIHLGLRLQRTLSIDLEIQEKLVCIGSVGS 599
Db 629 -----DRKLHDTFNIKROLVAIVRGVS 653
Oy 600 KTSLSIALIGQMTLLGSLAISGTFAYVAQOAMITNATLRDNIILGKEYDERIVSYLN 659
Db 654 GKSSLLHALIGEMNKLSGVSQVQVAYVOLAMIONLSIRNNILNRPYDALYONVIE 713
Oy 660 SCCLRPDLALPSSDLEIGERCANTSGGORISILARALYSRSYIILDDPALSADAHY 719
Db 714 NCALVODLESIPREDTEIGEKINISGGOKQVSLARAVYQNAELVLLDDPILSADSHY 773
Oy 720 GNIIFNSAIRKH--LKSRTVLEVTHQLOYVDCDEVIFMKEGCITERTGTHEELMNLND 776
Db 774 GKIIFFENVISATGCTGKTQVLLTGLTYLKHCDQVIVLTKDETISEMGTYQELMNSNGA 833
```

```

QY 777 YATIFNNLLGE-----TPPEINSKKETSGSKKSD 809
      ::::|
Db 834 FSEFLFEELLESKHKGRSVSFGEDSKVENELLRLDQVSPALRQILOSQSDIEKTD 893
QY 810 KGP---KTSIKKER-----AVKPEE-----GOLVLEEKGOG 839
      |::|::|
Db 894 KNEEIIINGLHKGDTAHSIGSKSEKESLLGALSPEKPEPPKOTKQIIEKAVETG 953
QY 840 SVPSVYGVYIOAAGPLA---FLVIMALFMLNVGSTAFSTWMLSTW-----IKOGSG 889
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 954 KYKREYVMSYFRAIGIKIALVFLVYVASSMLGV---FSNLYLAWMSDAKETAISLNG 1009
QY 890 NTVTRENNESVSDSMNDPMHOYASIALYSMAVMLILKARIGVNVKGLRASSRLHD 949
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1010 SSSST-----QIRLGIVAVLNGQATSVCAASTIHALGWCSRLIHA 1052
QY 950 ELFRILRSPKPFEDPTPTGRILNRFSGKMDVDVRLPFOEMFIQVNIYLFVCGMIA 1009
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1053 TLENIMRSPMAFEDVPLRGLILNRFGKVDVGDTPDIPRTMSMFIPTAVSSIEITAILM 1112
QY 1010 VEPFVLAVGPLVLSVLVSVLRVILREKRLDNTIOTSPPLSHITSSIOGLATIHAYNK 1069
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1113 ATPALISLLPLFRFYV-----STSRQIKRLSASRSPYTSRHOESIOGASSIRAVGV 1166
QY 1070 GQELHRYOELLDNOAPPELFTCAMRLAVRLDLI--SIALITTGIMIVLMHQ--IPPA 1127
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1167 VDKTRRSQHNVDNLATYPSIVANRLAVLEMGVLVLSAGAAVYRDSFGISAG 1226
QY 1128 YAGTAISYAVOLTGLFOFTVRLASETEARFYSVERINHYITLSLEAPRIKNAKPSDM 1187
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1227 LVGLSVSQAINTQITLMMAVAVMTSELETNIYAVRINEY--TITPTBGNSSQSLAPK--SW 1283
QY 1188 PQEEVYFENAEKMYRERPLVYLKKVSTFKPKKIGIVGTGSGKSSLGALRLVYLS 1247
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1284 PENEISIKNSVYRQGLDVLHGVTAHISPCERIGIVGTGAGKSSLTALRILEAD 1343
QY 1248 GGCIGKGVNRSDIGLADLRKSLIIPQEPVLFSGTVSNIDPFNOYEDQIMDALERTH 1307
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1344 GGCLEIDSTNADLLERLRLITIVODPVLFGSTAMNIDPFPAEDDQIWEILRAH 1403
QY 1308 MKECIADLPKLESEVMENGDNFSVEROLLICARALLRHCKIILDEATAMDETDL 1367
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1404 LDFSVKSLQEGILHNHISGEGNLSVGRQLICLALRLKTKVLVLEDAAAVDETDL 1463
QY 1368 IOETREAFACSTMLTNAHRLHYLGSRLVLAOGVVEEDTSSVLLSNDSSRYAM 1425
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1464 LQKTRIQFKDCTVLTAAHRLNTWMDSDRLVLDKGVAEEDTPKLLSNDGIFTSM 1521

RESULT 11
ID 094136 PRELIMINARY; PRT; 1540 AA.
AC 094136;
DT 01-FEB-1997 (TREMblrel. 02, Created)
DT 01-FEB-1997 (TREMblrel. 02, last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
DE MULTIDRUG RESISTANCE RELATED PROTEIN 1.
GN MRP-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Dean M., Gerrard B., Allikmets R., Broeks A., Plasterk R.H.A.;
RL EMBO J. 0:0-0(0).
CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
TRANSPORTERS).
CC EMBL; U66260; AAB07021.1; -.
DR HSSP; P13569; INBD.
DR InterPro; IPR001140; -.
DR InterPro; IPR001687; -.
DR InterPro; IPR003439; -.

```

```

DR InterPro; IPR003593; -.
DR Pfam; PF00005; ABC_tran; 2.
DR Pfam; PF00664; ABC_membrane; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
DR SMART; SM00382; AAA; 1.
KW ATP-binding; Transport.
SO SEQUENCE 1540 AA; 172410 MW; BA12FAD1341EBB9D CRC64;

Query Match 27.48; Score 2005; DB 5; Length 1540;
Best Local Similarity 33.68; Pred. No. 1.4e-127;
Matches 487; Conservative 240; Mismatches 483; Indels 240; Gaps 28;

QY 96 KHQPVNAGLFGSCFTSWLSLARVAHKKGELSMEDVWSISKHSSDVCRLERLMOE 155
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 204 KNSCPETASTLNDLTQWNSGLAYLGNK--SLKEDLDMLNED-----KRENIIPS 255
QY 156 ELNVEGPDASLR-----VWIFCRTRL--IISIVCL--MTOLACPSG 196
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 256 FIENLPEVEGRRKIKKNPEAIPKNHPSILIPFTYKFTLLAGCGYKLMEDLQFVA 315
QY 197 PAFVAKHLETTQATFESLQYSLVGLLTETVRSMSLATWALYRQVRLRGAILT 256
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 316 PE-LIRQLISFIEDKNQPMWIGVIALIMFLSSLLQSMILHQYHEMERLGMNRSVLS 374
QY 257 MAFKKILIKKN--IKESGLIELINICSDGORMFEAAVSGSLAGPVVALIGMIVYII 314
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 375 AVYTKTLNLSNARKKGTGAIVMLMSYDIDQDMTFTMLFMSAPLQILSLYFLMK 434
QY 315 LGPTGFLGSAVLELFYPMAMFASRLTAVERKCYAANDERVQKNEVLYIKFTKTAWV 374
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 435 LGVSLAGFVILLIIPNSFISVKMRNCQEMOKLXDERIKMASEILINGKVKLYSWE 494
QY 375 KAFSQSVQKTEEBERILKAGYQSTIV---GVAPIVYVLAIVVTSVHNTLGF--L 428
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 495 KSMKMYLEVERKEIRVKKLSTYNAATLTSWACAPFLV--AVLTGLVYLMDENNVL 551
QY 429 TAAQAFVTVVFNSTFALKTYPFSVKSLSSEASVAVDFKSLFLMEEVHMKKNRPSPHI 488
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 552 TPQITFVALFNILRFLAVFAVYFSAVQCSASNRLKKEFAAEEH-----SPQT 603
QY 489 K-----TEMKATLAWDSHSSIONSPEKLPKKMKDRASRKKERYQLQRTHEQ 539
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 604 SLAVGTDSAIKMDGSPAW-----GSKEE----- 628
QY 540 AYLAQKHLHLLSDERSPEEBEKGHLHGLRLQRTLSIDLEIQSGKLVGIGSVGS 599
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 629 -----DRKLHDTFENIKRGQLVALVGRVS 653
QY 600 GKTSLISAILGOMTLLEGSIASISGFAYVAQAMILNTALNDNLIFGKEYDERNSVLN 659
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 654 GKSSILHLLEGMKLSGVQVNGSVAVYLPOLAMIQNLSTNNILFNRPYAKILQVNI 713
QY 660 SCCLRPDLAIPSSDLTEIGRGANLGGQROKISLARALYSDRSIYLDDPLSALDAHV 719
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 714 NCALYQDLESIPAEDRTEITGKGINLSGQKORVSLAAVYQNEIYVLLDDPLSAVDHV 773
QY 720 GNHIFNSAIRKH---LAKTYLVETHQIYVDCDEVYFMKEGCTTEGTGTEELMNLNGD 776
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 774 GKHLFENYISTATGCLGKTFRVLLTHGLTYLKHCDQVVLKDEITSEMGTYOELMNSNG 833
QY 777 YATIFNNLLGE-----TPPEINSK 797
      ::::|
Db 834 FSEFLFEELLESKHKGRSVSFGEDSKVENELLRLDQVSPALRQI 893
QY 798 KETSGOKSODKGP---KTSIKKER-----AVKPEE----- 827
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 894 SOMSOEIKETDKNAELIRNGLHDEQTAHSIGSKSEKESLLGALSPEKPEPPKOTK 953
QY 828 GOLVLEEKGOGSVPWSVYGVYIOAAGPLA---FLVIMALFMLNVGSTAFSTWMLSTW- 883
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 954 TQLIKEAVEIGKVFYFVMSYFRAIGIKIALVFLVYVASSMLGV---FSNLYLAWMS 1009

```

QY 884 -----IKSGSNTVTYRGNETSVDSSKMDPNHMOYASIVALSMAVMLLKAINGVVFV 937
Db 1010 DDAKIALISGMSSET-----OIRLGITAVLGMGATSCAASITMA 1052
QY 938 KGTLRASSRLHDELFRRLILRSPMKFFDTPTGRLINRFSKDMDEVAVLRPEQAEFIONV 997
Db 1053 LGMVCAASHLHNTLEENRMSPMAFDVTPGLRLINRFGKDVGDVTDIPRMSFIRTA 1112
QY 998 ILVFCVGMAGVFPWFVLVAVGPVLVLSVLAIVSRVLIRELKLNDITQSPFLSHITSS 1057
Db 1113 VSSIRIILIIIMAPPLAISLLPLPIRFYV-----STRQIKRLESASRSPIYSHFQES 1166
QY 1058 IGGATTAAYKNGOELRHYOEILDNDQAPFLFTCAMRWLAVRDLI-STALITTTGLM 1116
Db 1167 IGGASIRAYGVNDKFIRESQHRVDENLATIYPSIVANRWLAVRLEMGNGLVLVSAGAA 1226
QY 1117 IYLMHQ-IPPAVAGLAISAVOLTGLEFQFVRLASETEARTSVERINHYIKTISLEAP 1175
Db 1227 VYFRDSPGLSAGLVGLSYVALNITQTLNMAVRMTSELETINVAVERINEX--TTPREG 1284
QY 1176 ARIKKASPDMPDGEVTFENAEKRENDLPVLTKVYSFTTKPEKIGIVRTSGSKSS 1235
Db 1285 NNSOLAER-SWPEGEISIKNFVRYRPGDLVLHGYTAHISPECKIGIVRTGAKSS 1343
QY 1236 LGMALFRLVELSGGCIKIDYARISDIGLADLRKSLITIPQEPVLFSGVYRSLDPEFNOYT 1295
Db 1344 LTLALFRILEADGGCIEIDGTNIADILLEQLRSRLTIVPQDVLFSGRKRMMLDPPAFS 1403
QY 1296 EDOIDALERTHMKECIALQPLKLESEVWENGDNFVSEROLLICARALLRCKILLIDE 1355
Db 1404 DDQIWEALRNHLDSEFVKSLOGELHHHISEGENDLSVGROQLICARALLRCKLVLIDE 1463
QY 1356 AFAANDTETDLIQTITFAPDCMTLITANRLHVLVLSDRIMVLAQGVYFEDIPSVLL 1415
Db 1464 AAAAVDVFETSLQGTIRQPKDCTVLTIAHRLNTVMDSRLVLIDKGVAFEDIPKAVL 1523
QY 1416 SNDSRFYAM 1425
Db 1524 SNPDGIFYSM 1533

RESULT 12
ID 048907 PRELIMINARY: PRT: 1623 AA.
AC 048907:
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE MGATP-ENERGIZED GLUTATHIONE S-CONJUGATE PUMP (EC 3.6.1.3).
GN MRP2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu Y.-P., Li Z.-S., Drozdowicz Y.M., Hortensteiner S., Martinoia E.,
Rea P.A.;
RC Plant Cell 10:1-18(1998).
CC -I- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
TRANSPORTERS).
CC EMBL: AF020288; AAC04245.1; -.
DR HSSP: P13569; INBD.
DR Mendel: 24660; Arabid.1944;24660.
DR InterPro: IPR001140; -.
DR InterPro: IPR001687; -.
DR InterPro: IPR003439; -.
DR InterPro: IPR003593; -.
DR Pfam: PF00005; ABC_tran. 2.
DR Pfam: PF00664; ABC_membrane. 2.
DR PROSITE: PS00211; ABC_TRANSPORTER. 2.
DR SMART: SM00382; AAA; 1.

KW ATP-binding; Hydrolyase; Transport.
SQ SEQUENCE 1623 AA; 182072 MW; 9630B7B04899970 CRC64;
Query Match 27.3%; Score 1993; DB 10; Length 1623;
Best Local Similarity 34.7%; Pred. No. 9,8e-127;
Matches 481; Conservative 230; Mismatches 483; Indels 192; Gaps 29;
QY 95 CKHQHPVDNAGLFSCMTFSWLSLARVAHKKGELSMEDVWSLSKHESSDYNCRLERLMQ 154
Db 226 CPEKHP-----NIFDKIFFSWMNPLMTLSKR-PLTEKVYVLDITWDQETLETSPQSHD 280
QY 155 EELNVPDPAASLRVWVVFCTRLILSLIVCLMTIQLAG-----FSGPA 198
Db 281 KELQKPP-----W-----LRLANSLSGRFMGGFWRKIGNCSQFVGPL 321
QY 199 F---VWKHLEYTOATESNL-OYSLL--LVGLLLETIVRSWLSALTYALNRYGVRLNG 252
Db 322 LNLQGLKSMQEDAPAMMGITYAFSLFEGVFEGLCE-----AQYFQNVKRVGRLRS 373
QY 233 ALLTMAFKKILKLNKIKESL--GELINICSDGQRMFEAAVGSLLAGPVAILGMIT 310
Db 374 ALIAAVFRKSLRLTEGRKKFPQTKITMTDASLQOICSLHTMMSAPRIIALL 433
QY 311 NVTLIGPGLGSAVFILFEPAMFASRLTAYFRKKVATDERYQKNMEVLTYIKFIK 370
Db 434 LYQGLVASLIGALLIVLMFPIQYIYSKQKLTKEGLQRTIKRIGLNEVLAAMDYKVC 493
QY 371 YAVMAFSGSVOKIREERIRILEKAGYFOSITVGVAVPIVAVTSVHNTLGFDLTA 430
Db 494 YAMENSFQSKVQTVRDELDLSWRKSQLGALNMFILNSIPVLVTIVSFQVFLTGLDITP 553
QY 431 AQAFTVYVFNKMTALAKYTPFSVKSLSASVAVDRFKSLFMEVYHMKNNKPNAP-HIK 489
Db 554 ARAFTSLSLFAVLRPPLFPLPNIITQVNAVANSLNRLKEVLAETERILLPNPIPEGPA 613
QY 490 IEMKNATLAMDSSHSSIONSPLTPPKMKDKRASRGKKEKVQLORTEHQVLAQKHL 549
Db 614 ISTRNGYFSWDS-----KGS-- 627
QY 550 LLDSDERSPEEKEGKHNIHGLRLQRTLHSLDLEIOEGKLVGSGVSGKTSLSIALT 609
Db 628 -----DRP-----TLSNINLDVPLGSLVAVVSGTEGTSLSIALT 663
QY 610 GOMTLLEGSI-AISGFAYVAQOAMTLNATLNDNLIFGKYDEERYNSVLNSCCLRPDLA 668
Db 664 GELPRTSDIVTLRKSVAVVPQVSWIFENATVNDNLIFGSPFPRKEXERAIDVTSKHDLE 723
QY 669 ILPSSDLFPIEGRGANLGGGQROKRSISARALYSDRSIYLLDDPLSLADHVGNIHFNSAI 728
Db 724 LIPGDULTEIGERGYNISGGQKORVSMARAVYSNSDVYLFDDPLSLADHVGQOYFEKCI 783
QY 729 RKLHLSKTYLFTYTHLOLYLVDCDEVYFMKEGCTERGHHEELMNINDYATLFNNLLIG- 787
Db 784 KRELQKTFVLVTLNQHFLFSQYDRIVLVHEGTVEKEGYTEELSS--NG--PLFQRLMEVA 839
QY 788 -----ETPPEVINSKKE-----TSGSQK--SQDKPRTGSITKEKAVKPEBGQLV 831
Db 840 GKYEEYSENGEAEADQTEQPVANGNTGLMDSDDKSEKNGKKGKSV-----LI 893
QY 832 QLEEKGGGVPSYSGYVYQAAGGPLALVIALFPLANGSGTAFTSMVLSYVIKQSGMT 891
Db 894 KOBERETGVASRWVLLKRYODALGGAMVMMLLCYLLEVFVRYTSTWLSMTWTDAG--T 950
QY 892 YTVRGNETSVDSSKMDPNHMOYASIVALSMAVMLLKAIRVAVFVFKGLRASSRLHDEL 951
Db 951 PKSHG-----PLFYNLITALLSFGQYLVLTITNSYWLIMSLVYAAKKLHDMA 996
QY 952 FRRLILRSPMKFFDTPTGRLINRFSKDMDEVAVLRPEQAEFIONVILVFCVGMITAGVF 1011
Db 997 LHSILRAPMSPFHTNPLGRILINRFAKDLGIDIDRTVAVFVNMFGVQSOLSTVVLGIYS 1056
QY 1012 PMFLVAVGVLVLSVLAIVSRVLIRELKLNDITQSPFLSHITSSIGSLATIIHAYNKQ 1071


```

Db 1366 VPDPVLFSGTMKMLDPEFSAVSDSQWEALENAHLKRFVSLDQGLEHKISKGGENLSV 1425

Oy 1333 GEROLLCIARALLRHCKLLIDEATAMODETDLIQETIREAFADCTMLTAHRLHVL 1392
Db 1426 GOROLICARALLRKRTKVLVDEAAAVADVDSITOKTIREOFECFCTVTAHRLNTVM 1485

Oy 1393 GSDRIMVLAGOVVEPDTSPVLSNDSSRFYAM 1425
Db 1486 DSDRLVLVDKGVAEFSDSPKNLLANPDGIFYSM 1518

RESULT 15
020943
ID 020943 PRELIMINARY: PRT: 1494 AA.
AC 020943;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE SIMILAR TO THE HUMAN MULTIDRUG RESISTANCE-ASSOCIATED PROTEIN.
GN F57C12.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnscoough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copest T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jler M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latelle P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Woildman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX
RA Favello T.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX
RA Waterston R.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
CC TRANSPORTERS).
DR EMBL; U41554; AA83299.1; -.
DR HSSD; P13569; INMD.
DR InterPro; IPR001005; -.
DR InterPro; IPR001140; -.
DR InterPro; IPR001687; -.
DR InterPro; IPR003439; -.
DR InterPro; IPR003593; -.
DR Pfam; PF00005; ABC_tran; 2.
DR Pfam; PF00664; ABC_membrane; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
DR PROSITE; PS00037; MIB_1; UNKNOWN_1.
DR SMART; SM00382; AAA; 1.
KW ATP-binding; Transport.
SQ SEQUENCE 1494 AA; 16742 MW; A88A6EC8F683AFDC CRC64;

Query Match 27.2%; Score 1984.5; DB 5; Length 1494;
Best Local Similarity 34.5%; Pred. No. 3.2e-126;
Matches 493; Conservative 232; Mismatches 499; Indels 203; Gaps 31.

```

QY	152	LMQBELNVEYGDASLR-----VWVF--CRRLILSIYCMITDLAGSSPAFVWHL	205
Db	222	YH0IKKD--PSAALPKNHPSPVIFPKYKYTLLAGFFYKLCEDMLQLAQ--LTKLI	278
QY	206	EYTOATESNIQYSLILVGLLLEIYRSWSLALTMALNRTGVRLGALITMAEFKIL	265
Db	279	GFIEDKNQPMWIGSIVGIMFSSPLOSMTLHQYHSMRFLGMHVRSLVTSKALNL	338
QY	266	KN--IKESJIGELINICSNDGQMFPAAVGSLLAGPVALGMIVNIIIGPFGIS	323
Db	339	SNEARKKTTGAIYNLMSVDIOKIDOMAPITMLFMSAPLOIFLSTYFMKEFGAVALGL	398
QY	324	AVFLTEPAMMFASRLAYPRRC---VAATEDROKNEVLEITFKIFMAYAWAFSQ	379
Db	399	VVLLALP-----VGLALIDMRKCOQEDOMKLDERIKMSELINGKYLKXSWRSMEN	454
QY	380	SVOKIREEREERILEKAGYFOSITGVAPIVVIVIASVTEVSMTLGF--LFAAQFTV	436
Db	455	MLKIRREELHILIKLSYFMAAIVFESWICAPLAIISFVVVVYVUDPENNVLTPEITFA	514
QY	437	VTVNSMTFALKVTPESVSKLSASAVVDRFSLFLMEVNHKKKPSPHKITEMKAT	496
Db	515	LLEPDLIRMLPAMAYVGEAVOCVSANRLEKFAEEM-----SPQSTI-----	560
QY	497	LAMPSSHSI0NSPKFLTPKKKKKKRKRSGKKEKVR0LQTEHQAVLAEQGLHLDSDER	556
Db	561	-----SHG-----EIDSAIIEVENGFLSMSDED	583
QY	557	PSPEEBEKKHILHGLRLQRTLHSIDLEIOEGLVIGCSVSGKTSLSIALGOMTLE	616
Db	584	P-----TLREISFKI0KG0LAVIKVSGSSLEHLAGEMNLS	624
QY	617	GSIAISGFAVAAQAAWILNATLROMLTEKEDERYNSVNSCCLRPDLALPSSDT	676
Db	625	GSVQINGRIAYVQOAMITOMSLRNMLFKPDLLENEDYKKNCAKLEDIANLPAGRT	684
QY	677	EIGERGANLSCG0R0RISARALYSDRSYIIDDPLSALDAHVGNHIFNSAIRKH--Lk	733
Db	685	EIGERGINLSSGQK0RYSLARAYOINDDIILDDPLSADSVHGKHIFENVISSTGCLA	744
QY	734	SKTYLPTFHOLOYVDCDEVIFPKKEGICTERGHEHMLNDGVATFINLL-----	785
Db	745	SKTYLVTHGLTYLKHCDOLIVKEGTISELGTQY0LLNNSCARFLEELFLESKTRG	804
QY	786	-----LQETPP-----VEINSKKTSG-----SOKKS--	807
Db	805	RVASIGDOSGEVDLRLDQVKKPGLIKRLLESHSQESKEEDTSARATIEYSDSSRRSVL	864
QY	808	-----Q0KGKPTGSIKKKEKAVKBEQOVLQLEKGGSPWGVYGYIQAAGPLAF	859
Db	865	LHSPRSGHEEHALGAIISDVPDQ--BNOLLEKEKETVETGKKFEFYIAYFALISIPTL	923
QY	860	LVIALLFMLNAGSTAFSTWMLSYWIKOGSQNTVTGNETSYSDSKNDPNHMOYASIIYA	919
Db	924	L-----PFLYVSGSGJGILNFYLAFL-----SDHAKSGNNTSDAKME-----LGIYA	968
QY	920	LMAVMLILKAINGVAVFKCTLAASSRLDELFRLLRSLPKMFPTTTPGRLIRFSQDM	979
Db	969	VLCMG0SFVYLIIISLITIVLRASRLIAGLGNIMRSPMAFPVPTPIGRILNIGKDI	1028
QY	980	DEV0VRLPFOAEMFI0WLVLFPCVMINGVFPWFVVAAGPVLIEFSVLHYSRVL---	1036
Db	1029	EADIRLTPDVIIRHSMITINNVATLVYIMATPMAIGIA---FALLSVIYFILPREYISTS	1085
QY	1037	RELKRLNITQSPFLSIITSSIOGLATTHAYNKG0EELHRY0ELLDN0AQPFLETCMR	1096
Db	1086	ROLKRLSASRSPISYF0ESIOGASSIRAFGVDFNFKOS00RVDHILIAVYSIVANR	1145
QY	1097	MLAVRDLI--SIALITTYGMIYLMHGQ--IPRAYGILASVVOULTGLFQFVPLASPTE	1154
Db	1146	WLAVRELVGVLVLSAAGAVYFRDSPGSLAGVLGSLVSTALNTI0PLNMAVVRTSLE	1205

Tue Nov 13 11:55:03 2001

us-09-528-031-2.rsp

Page 17

```

QY 1155 ARFVSYRIHHYKTTLSLEPARIKK--KAPSPMDQCEGTEFNAEKRYEENPLVYKK 1212
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1206 TNYVSERIEHYVT-----PFBGNNSRRLAAKSNMPEKGEISIKNFSTRYRPGDLVLHG 1260

QY 1213 VSFYIKPEKIGIVGTSGKSSGLNALFRVLEYSGCCIKIDGVIRISDIGIADIRSKLSI 1272
      :|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1261 ISAHIAPESEVGVIGVGTGAKSKSLTALFRILIEADGGSIEIDGIAMIAOLEPQRSCTII 1320

QY 1273 IPOEPLYFSGTGVSRNLDPPNQTAEODIMPALEPRTMKECIQQLPLKSEWEKND--- 1328
      :|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1321 VPDDPLYFSGTGMKMNLDPPSATSSDQSWPALENAHUKRPVKSLODGLLHKKISSEGENURF 1380

QY 1329 -----NFSVGEROLLICARALRLRRCKTLLIDEATAAMDTEFDLLIQETIREAFAD 1378
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1381 AGSKSKQFOELNFSVSGORQICARALRLRRCKTLYLVDEAAAAVDVETDLSIQKTIREQFKE 1440

QY 1379 CTMKTALHRLHVTLGSDRIIMVLAAOGVSEFDFPVSULLINDSSRFPAM 1425
      ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1441 CTYVITIAHRLNTYMDSDRLVLVDKRGVAEFDSPKMLINLAPDQICFYSM 1487

```

Search completed: November 9, 2001, 18:04:44
Job time: 1077 sec

THIS PAGE BLANK (USPTO)